

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 01:14:10 ; Search time 132 Seconds
(without alignments)
1115.090 Million cell updates/sec

Title: US-10-733-782-1
Perfect score: 1678
Sequence: 1 AIDENKQKALAAALQIEKQ.....AKIEKKVRELLGNPNSTP 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: Geneseqp1980a:*
2: Geneseqp1990a:*
3: Geneseqp2000a:*
4: Geneseqp2001a:*
5: Geneseqp2002a:*
6: Geneseqp2003a:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1678	100.0	335	8	ADR16428	Adr16428 E. coli R
2	1678	100.0	352	8	ADR16432	Adr16432 E. coli w
3	1678	100.0	353	6	ABU15395	Abu15395 Protein e
4	1678	100.0	358	3	AY68827	AY68827 Amino aci
5	1678	100.0	358	3	AY68829	AY68829 Amino aci
6	1678	100.0	358	8	ADQ87824	Adq87824 E. coli r
7	1678	100.0	358	8	ADQ87822	Adq87822 E. coli r
8	1678	100.0	358	8	ADQ87821	Adq87821 E. coli r
9	1677	99.9	358	3	AY68831	AY68831 Amino aci
10	1674	99.8	335	8	ADR16430	Adr16430 E. coli R
11	1674	99.8	358	2	AAW64213	AAW64213 New minsh
12	1673	99.7	358	3	AY68828	AY68828 Amino aci
13	1673	99.7	358	8	ADQ87823	Adq87823 E. coli r
14	1669	99.5	358	8	ADQ87826	Adq87826 E. coli r
15	1668	99.4	358	2	AAW64215	AAW64215 Hyperreco
16	1665	99.2	358	2	AAW64214	AAW64214 Hyperreco
17	1662	99.0	358	2	AAW64217	AAW64217 Hyperreco
18	1662	99.0	358	2	AAW64218	AAW64218 Hyperreco
19	1660	98.9	358	3	AY68832	AY68832 Amino aci
20	1660	98.9	358	8	ADQ87827	Adq87827 E. coli r
21	1659	98.9	358	3	AY68830	AY68830 Amino aci
22	1659	98.9	358	8	ADQ87825	Adq87825 E. coli r
23	1643	97.9	358	2	AAW64216	AAW64216 Hyperreco
24	1641	97.8	353	6	ABU48024	Abu48024 Protein e

25	1613	96.1	352	6	ABU31601	Abu31601 Protein e
26	1613	96.1	356	7	ABO63433	ABO63433 Klebsiell
27	1593	94.9	352	6	ABU27708	Abu27708 Protein e
28	1547	92.2	356	6	ABM67793	ABM67793 Phototrab
29	1534	91.4	356	6	ABU50600	ABU50600 Protein e
30	1504	89.6	355	6	ABU40869	ABU40869 Protein e
31	1504	89.6	367	7	ADF05150	Adf05150 Bacteriol
32	1467	87.4	351	6	ABU44971	Abu44971 Protein e
33	1436	85.6	412	6	ABU49151	Abu49151 Protein e
34	1353	80.6	354	6	ABU39491	Abu39491 Protein e
35	1332	79.4	354	6	ABU30311	Abu30311 Protein e
36	1269	75.6	398	7	ABO72381	ABO72381 Pseudomon
37	1268	75.6	346	6	ABU38581	Abu38581 Protein e
38	1259	75.0	349	6	ABU16844	Abu16844 Protein e
39	1259	75.0	355	6	ADA36276	Ada36276 Acinetoba
40	1257	74.9	355	6	ABU40018	Abu40018 Protein e
41	1255	74.8	369	6	ABU41878	Abu41878 Protein e
42	1250.5	74.5	366	8	ADL05025	Adl05025 M. catarr
43	1245.5	74.2	349	6	ABU35468	Abu35468 Protein e
44	1241.5	74.0	348	6	ABU33125	Abu33125 Protein e
45	1241.5	74.0	348	9	ABE39711	Aeb39711 L. pneumo

ALIGNMENTS

RESULT 1
ADR16428
ID ADR16428 standard; protein; 335 AA.
XX
AC ADR16428;
XX
DT 04-NOV-2004 (first entry)
XX
DE E. coli RecA mutant protein #1.
XX
KW RecA; DNA binding protein; ssDNA binding protein; SSB;
KW DNA strand exchange reaction; site-specific cleavage; mutant; muten.
XX
OS Escherichia coli.
XX
PN US2004157248-Al.
XX
PD 12-AUG-2004.
XX
PF 11-DEC-2003; 2003US-00733782.
XX
PR 12-DEC-2002; 2002US-0432758P.
XX
(COXM/) COX M M.
PA (LUSE/) LUSETTI S L.
PA (EGGL/) EGGLER A L.
XX (HARU/) HARUTA N.
XX
PI Cox MM, Lusettti SL, Eggler AL, Haruta N;
DR WPI; 2004-580265/56.
DR N-PSDB; ADR16429.
XX
PT New RecA mutant proteins comprising a single mutation or a double
PT mutation, useful for catalyzing homologous DNA pairing and DNA strand
PT exchange reactions in an in vitro or in vivo environment.
XX
PS Claim 2; SEQ ID NO 1; 34pp; English.
XX
CC The invention relates to an isolated RecA mutant protein, where the
CC protein is either a single mutant RecA protein comprising a deletion of
CC amino acid residues from the carboxyl terminus, or a double mutant RecA
CC protein comprising a deletion of amino acid residues from the carboxyl
CC terminus and an amino acid change from a glutamate to a basic amino acid.
CC The invention also relates to a polynucleotide sequence that encodes the
CC RecA mutant protein, a method of catalysing an in vitro homologous DNA
CC pairing and DNA strand exchange reactions comprising providing an amount

CC of the RecA mutant protein and a method of increasing recombination
 CC efficiency of homologous DNA pairing and DNA strand exchange reactions in
 CC a cell comprising supplying to the cell an amount of the RecA mutant
 CC protein. The RecA mutant protein comprises an enhanced capacity to
 CC displace a DNA binding protein as compared to wild-type RecA. The DNA
 CC binding protein is the ssDNA binding protein from *Escherichia coli* (SSB).
 CC It also comprises enhanced binding to DNA during a DNA strand exchange
 CC reaction as compared to wild-type RecA. The composition and methods are
 CC useful for catalyzing homologous DNA pairing and DNA strand exchange
 CC reactions in an in vitro or in vivo environment. These may be used in
 CC promoting in vitro alterations of genes to permit the rapid construction
 CC of desired gene mutants for industrial and pharmaceutical purposes. The
 CC mutant proteins may also be used in targeting the site-specific cleavage
 CC of small and large DNAs, or as a basis for the design and construction of
 CC tiny electronic circuits based on DNA. This sequence represents a RecA
 CC mutant protein of the invention.

XX Sequence 335 AA;

Query Match 100.0%; Score 1678; DB 8; Length 335;
 Best Local Similarity 100.0%; Pred. No. 5.7e-155;
 Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSDVETISTGSLDLTALGAGGLPMGR 60
 DB 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSDVETISTGSLDLTALGAGGLPMGR 60

QY 61 IVEIYGPSSSGKTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDIDNLLCSQPD 120
 DB 61 IVEIYGPSSSGKTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDIDNLLCSQPD 120

QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGIAARMMSQAMRKLAG 180
 DB 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGIAARMMSQAMRKLAG 180

QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVGENVVG 240
 DB 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVGENVVG 240

QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLIKAGAWYSYKGEKIQ 300
 DB 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLIKAGAWYSYKGEKIQ 300

QY 301 GKANATAWLKONPETAKEIEKKVRELLSNPNSTP 335
 DB 301 GKANATAWLKONPETAKEIEKKVRELLSNPNSTP 335

RESULT 2
 ID ADR16432
 AD R16432 standard; protein; 352 AA.

XX AC ADR16432;
 XX DT 04-NOV-2004 (first entry)
 XX DE *E. coli* wild-type RecA protein.
 XX KW RecA; DNA binding protein; ssDNA binding protein; SSB;
 KW DNA strand exchange reaction; site-specific cleavage.
 XX OS *Escherichia coli*.
 XX PN US2004157248-A1.
 XX PD 12-AUG-2004.
 XX PF 11-DEC-2003; 2003US-00733782.
 XX PR 12-DEC-2002; 2002US-0432758P.
 XX PA (COXM/) COX M M.
 PA (LUSE/) LUSETTI S L.

PA (EGGL/) EGGLER A L.
 PA (HARU/) HARUTA N.
 XX Cox MM, Luseti SL, Egglar AL, Haruta N;
 XX WPI; 2004-580265/56.
 XX New RecA mutant proteins comprising a single mutation or a double
 PT mutation, useful for catalyzing homologous DNA pairing and DNA strand
 PT exchange reactions in an in vitro or in vivo environment.
 XX Disclosure; SEQ ID NO 5; 34pp; English.

The invention relates to an isolated RecA mutant protein, where the
 CC protein is either a single mutant RecA protein comprising a deletion of
 CC amino acid residues from the carboxyl terminus, or a double mutant RecA
 CC protein comprising a deletion of amino acid residues from the carboxyl
 CC terminus and an amino acid change from a glutamate to a basic amino acid.
 CC The invention also relates to a polynucleotide sequence that encodes the
 CC RecA mutant protein, a method of catalyzing an in vitro homologous DNA
 CC pairing and DNA strand exchange reactions comprising providing an amount
 CC of the RecA mutant protein and a method of increasing recombination
 CC efficiency of homologous DNA pairing and DNA strand exchange reactions in
 CC a cell comprising supplying to the cell an amount of the RecA mutant
 CC protein. The RecA mutant protein comprises an enhanced capacity to
 CC displace a DNA binding protein as compared to wild-type RecA. The DNA
 CC binding protein is the ssDNA binding protein from *Escherichia coli* (SSB).
 CC It also comprises enhanced binding to DNA during a DNA strand exchange
 CC reaction as compared to wild-type RecA. The composition and methods are
 CC useful for catalyzing homologous DNA pairing and DNA strand exchange
 CC reactions in an in vitro or in vivo environment. These may be used in
 CC promoting in vitro alterations of genes to permit the rapid construction
 CC of desired gene mutants for industrial and pharmaceutical purposes. The
 CC mutant proteins may also be used in targeting the site-specific cleavage
 CC of small and large DNAs, or as a basis for the design and construction of
 CC tiny electronic circuits based on DNA. This sequence represents the wild-
 CC type RecA protein of the invention.

XX Sequence 352 AA;

Query Match 100.0%; Score 1678; DB 8; Length 352;
 Best Local Similarity 100.0%; Pred. No. 6.2e-155;
 Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSDVETISTGSLDLTALGAGGLPMGR 60
 DB 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSDVETISTGSLDLTALGAGGLPMGR 60

QY 61 IVEIYGPSSSGKTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDIDNLLCSQPD 120
 DB 61 IVEIYGPSSSGKTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDIDNLLCSQPD 120

QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGIAARMMSQAMRKLAG 180
 DB 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGIAARMMSQAMRKLAG 180

QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVGENVVG 240
 DB 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVGENVVG 240

QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLIKAGAWYSYKGEKIQ 300
 DB 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLIKAGAWYSYKGEKIQ 300

QY 301 GKANATAWLKONPETAKEIEKKVRELLSNPNSTP 335
 DB 301 GKANATAWLKONPETAKEIEKKVRELLSNPNSTP 335

RESULT 3
 ID ABU15395
 XX ABU15395 standard; protein; 353 AA.

AC ABUL5395;
XX 19-JUN-2003 (first entry)
XX Protein encoded by Prokaryotic essential gene #922.
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Escherichia coli.
XX WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX N-PSDB; ACA19265.
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 43319; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 353 AA;
Query Match 100.0%; Score 1678; DB 6; Length 353;
Best Local Similarity 100.0%; Pred. No. 6.2e-155;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKOKALAAALQIEKQFGKSGIMRLGEDRSMDETISTGSLSLDIALCAGGLPMGR 60
DB 2 AIDENKOKALAAALQIEKQFGKSGIMRLGEDRSMDETISTGSLSLDIALCAGGLPMGR 61
QY 61 IVEIYGPSSSGKTTLTQVIAAAQREGTKCAPIDAEHALDPIYARKLGVDINLLCSQPD 120
DB 62 IVEIYGPSSSGKTTLTQVIAAAQREGTKCAPIDAEHALDPIYARKLGVDINLLCSQPD 121
QY 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEITEGISHMGLAARMMSOAMRKLKAG 180
DB 122 TGEQALEICDALARSGAVDVIVDSVAALTPKAEITEGISHMGLAARMMSOAMRKLKAG 181
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGCGNALKFYASVRLDRIIRRGAVKEGNNVGS 240
DB 182 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGCGNALKFYASVRLDRIIRRGAVKEGNNVGS 241
QY 241 ETRVKVVKNTIAAPFKQAEFQILYGEINFGYELVDLGKVKELIEKAGAWSYKGEKIGQ 300
DB 242 ETRVKVVKNTIAAPFKQAEFQILYGEINFGYELVDLGKVKELIEKAGAWSYKGEKIGQ 301
QY 301 GKANATLWLDKNDPETAKEIEKKVRELLLSNPNSTP 335
DB 302 GKANATLWLDKNDPETAKEIEKKVRELLLSNPNSTP 336

RESULT 4

AAV68827
ID AAV68827 standard; protein; 358 AA.
XX AC AAV68827;
XX DT 16-MAY-2000 (first entry)
XX DE Amino acid sequence of a wildtype recA protein.
XX KW recA; hyperrecombinogenic variant; male gamete; female gamete;
KW heat tolerance; ethanol production; ethanol tolerance; metabolite; taxol;
KW cyclosporin A; erythromycin; meiosis.
XX OS Escherichia coli.
XX FN WO200004190-A1.
XX PD 27-JAN-2000.
XX PF 15-JUL-1999; 99WO-US015972.
XX PR 15-JUL-1998; 98US-00116188.
XX PA (MAXY-) MAXYGEN INC.
XX PI Del Cardayre S, Tobin M, Stemmer WPC, Ness JE, Minshull J;
PI Patten PA, Subramanian V, Castle LA, Krebber CM, Bass S, Zhang Y;
PI Cox T, Huiegan G, Yuan L, Affholter JA;
DR WPI: 2000-182446/16.
DR N-PSDB; AAZ60609.
PT Evolution of whole cells and organism by iterative cycles of
PT recombination and selection and screening for acquisition of desired
PT properties.
XX Example 1; Fig 13; 197pp; English.
XX CC The present sequence represents the recA protein. The wild type recA
CC protein was used to produce hyperrecombinogenic variants (see AAY68828-
CC 32), using the method of the invention. The specification describes a
CC method for producing a library of diverse multicellular organisms using
CC pools of male and female gametes. At least one of the male pool or female
CC pools comprises a number of different gametes derived from different
CC strains of a species or of a different species. The viable organisms
CC produced from the fertilized gametes are repeatedly crossed to produce a

CC library of diverse organisms, which are selected for a desired trait or
CC property. The methods can be used to evolve cells to acquire a desired
CC property such as heat tolerance, ethanol production or tolerance, acid,
CC improved production and maintenance of enzyme cofactors or NAD(P)H and
CC improved glucose transport. The desired property may be expression of a
CC protein or primary or secondary metabolite. Alternatively the desired
CC property is secretion of a protein or secondary metabolite, chosen from
CC taxol, cyclosporin A and erythromycin. The desired property may be a
CC capacity for meiosis or compatibility to form a heterokaryon with another
CC strain
XX
SQ Sequence 358 AA;

Query Match 100.0%; Score 1678; DB 3; Length 358;
Best Local Similarity 100.0%; Pred. No. 6.3e-155;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLDLALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLDLALGAGGLPMGR 66
QY 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGLAARMMSQAMRKLAG 180
DB 127 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGLAARMMSQAMRKLAG 186
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVGS 240
DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVGS 246
QY 241 ETRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLEKAGAWYSYKGEKIQ 300
DB 247 ETRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLEKAGAWYSYKGEKIQ 306
QY 301 GKANATAMLKONPETAKETIEKKVRELLSNPNSTP 335
DB 307 GKANATAMLKONPETAKETIEKKVRELLSNPNSTP 341

RESULT 5
AAY68829
ID AAY68829 standard; protein; 358 AA.
XX
AC AAY68829;
XX
DT 16-MAY-2000 (first entry)
XX
DE Amino acid sequence of a hyperrecombinogenic recA protein clone 4.
XX
KW recA; hyperrecombinogenic variant; male gamete; female gamete;
KW heat tolerance; ethanol production; ethanol tolerance; metabolite; taxol;
KW cyclosporin A; erythromycin; meiosis.
XX
OS Synthetic.
OS Escherichia coli.
XX
PN W0200004190-A1.
XX
PD 27-JAN-2000.
XX
PF 15-JUL-1999; 99WO-US015972.
XX
PR 15-JUL-1998; 98US-00116188.
XX
PA (MAXY-) MAXYGEN INC.
XX
PI Del Cardayre S, Tobin M, Stemmer WPC, Ness JE, Minshull J;
PI Patten PA, Subramanian V, Castle LA, Krebber CM, Baas S, Zhang Y;
PI Cox T, Huismen G, Yuan L, Affholter JA;
XX

DR WPI; 2000-182446/16.
DR N-PSDB; AAZ60611.
XX
PT Evolution of whole cells and organism by iterative cycles of
PT recombination and selection and screening for acquisition of desired
PT properties.
XX
PS Example 1; Fig 13; 197pp; English.
XX

CC The present sequence represents a hyperrecombinogenic recA protein. The
CC wild type recA protein (AAY68827) was used to produce hyperrecombinogenic
CC variants, using the method of the invention. The specification describes
CC a method for producing a library of diverse multicellular organisms using
CC pools of male and female gametes. At least one of the male pool or female
CC pools comprises a number of different species. The viable organisms
CC produced from the fertilized gametes are repeatedly crossed to produce a
CC library of diverse organisms, which are selected for a desired trait or
CC property. The methods can be used to evolve cells to acquire a desired
CC property such as heat tolerance, ethanol production or tolerance, acid,
CC improved production and maintenance of enzyme cofactors or NAD(P)H and
CC improved glucose transport. The desired property may be expression of a
CC protein or primary or secondary metabolite. Alternatively the desired
CC property is secretion of a protein or secondary metabolite, chosen from
CC taxol, cyclosporin A and erythromycin. The desired property may be a
CC capacity for meiosis or compatibility to form a heterokaryon with another
CC strain
XX
SQ Sequence 358 AA;

Query Match 100.0%; Score 1678; DB 3; Length 358;
Best Local Similarity 100.0%; Pred. No. 6.3e-155;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLDLALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLDLALGAGGLPMGR 66
QY 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGLAARMMSQAMRKLAG 180
DB 127 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGLAARMMSQAMRKLAG 186
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVGS 240
DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVGS 246
QY 241 ETRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLEKAGAWYSYKGEKIQ 300
DB 247 ETRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLEKAGAWYSYKGEKIQ 306
QY 301 GKANATAMLKONPETAKETIEKKVRELLSNPNSTP 335
DB 307 GKANATAMLKONPETAKETIEKKVRELLSNPNSTP 341

RESULT 6
ADQ87824
ID ADQ87824 standard; protein; 358 AA.
XX
AC ADQ87824;
XX
DT 09-SEP-2004 (first entry)
XX
XX E. coli recA protein clone #3.
XX
KW RecA; protoplast formation; gamete; heat tolerance; ethanol production;
KW ethanol tolerance; enzyme cofactor; NAD(P)H; glucose transport; meiosis;
KW heterokaryon; taxol; cyclosporin A; erythromycin; viral infection;
KW plant genome shuffling; microprobe manipulation; reiterative pooling.

XX OS Escherichia coli.
 XX AU2004200501-A1.
 XX PD 04-MAR-2004.
 XX PF 09-FEB-2004; 2004AU-00200501.
 XX PR 09-FEB-2004; 2004AU-00200501.
 XX PA (MAXY-) MAXYGEN INC.
 XX PI Yuan L, Cox T, Bass S, Castle LA, Patten PA, Ness JE, Tobin M;
 PI Affholter JA, Huismann G, Zhang Y, Krebber CM, Subramanian V;
 PI Minshull J, Stemmer WPC, Del Cardayre S;
 DR WPI; 2004-507924/49.
 DR N-PSDB; ADQ87817.
 XX

PT Evolving cells to acquire a desired property, by forming protoplasts of
 PT different cells, fusing protoplasts to form hybrid protoplasts (HP),
 PT producing regenerated cells, forming additional HP and producing
 PT additional regenerated cells.
 XX

PS Disclosure; Fig 13; 196pp; English.
 XX

CC The invention relates to a method of evolving cells to acquire a desired
 CC property by forming protoplasts of different cells, fusing the
 CC protoplasts to form hybrid protoplasts, incubating the hybrid protoplasts
 CC to produce regenerated cells, repeatedly forming protoplasts from
 CC regenerated cells, fusing protoplasts to form hybrid protoplasts in which
 CC genomes from protoplasts recombine to form additional hybrid genomes and
 CC incubating additional hybrid protoplasts for producing additional
 CC regenerated cells. The invention also relates to a method of producing a
 CC library of diverse multicellular organisms involving providing a pool of
 CC male gametes and a pool of female gametes, where one of the male pool or
 CC the female pool comprises several different gametes derived from
 CC different strains of a species or different species and the male gametes
 CC fertilise the female gametes, permitting at least a portion of the
 CC resulting fertilised gametes to grow into reproductively viable
 CC organisms, repeatedly crossing the reproductively viable organisms to
 CC produce a library of diverse organisms and selecting the library for a
 CC desired trait or property. The methods are useful for evolving cells to
 CC acquire desired properties such as heat tolerance, ethanol production,
 CC ethanol tolerance, improved production and maintenance of enzyme
 CC cofactors, improved production and maintenance of NAD(P)H and improved
 CC glucose transport. The desired property is the expression of a protein,
 CC primary metabolite or secondary metabolite, the secretion of a protein or
 CC secondary metabolite, capacity for meiosis or compatibility to form a
 CC heterokaryon with another strain. The secondary metabolite is chosen from
 CC taxol, cyclosporin A and erythromycin. The method is also useful for
 CC acquisition of desired properties such as enhanced recombinationicity,
 CC gene copy number or gene reductivity and capacity for expression and/or
 CC secretion of proteins or secondary metabolites. The method is further
 CC useful for predicting efficacy of a drug in treating viral infection,
 CC plant genome shuffling, micropore manipulation, producing transgenic
 CC animals, improvement of overexpressed genes for a desired phenotype and
 CC reiterative pooling and breeding of higher organisms. The cells evolved
 CC by the method are useful in molecular genetics. This sequence represents
 CC an E. coli recA protein clone used in the method of the invention.
 XX

SQ Sequence 358 AA;
 Query Match 100.08; Score 1678; DB 8; Length 358;
 Best Local Similarity 100.08; Fred. No. 6.3e-155;
 Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKSIIMRLGDRSDMVETISTGSLDIALGAGLPMGR 60
 DB 7 AIDENKQKALAAALGQIEKQFGKSIIMRLGDRSDMVETISTGSLDIALGAGLPMGR 66
 QY 61 IVEIYGPSSGKTTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDINDLLCSQPD 120

Db 67 IVEIYGPSSGKTTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDINDLLCSQPD 126
 QY 121 TGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGEIGDSHMGIAARMMSQAMRKLKAG 180
 Db 127 TGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGEIGDSHMGIAARMMSQAMRKLKAG 186
 QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVUGS 240
 Db 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVUGS 246
 QY 241 ETRVKVWNKTAAPFKQAEFQILYGEINFGELVDLGVKEKLIKAGAWSYKGEKIQG 300
 Db 247 ETRVKVWNKTAAPFKQAEFQILYGEINFGELVDLGVKEKLIKAGAWSYKGEKIQG 306
 QY 301 GKANATWLKONPETAKEIEKKVRELLSNPNSTP 335
 Db 307 GKANATWLKONPETAKEIEKKVRELLSNPNSTP 341

RESULT 7

ADQ87822
 ID ADQ87822 standard; protein; 358 AA.

XX AC ADQ87822;

XX DT 09-SEP-2004 (first entry)

XX DE E. coli recA protein clone #1.

XX KW RecA; protoplast formation; gamete; heat tolerance; ethanol production;
 KW ethanol tolerance; enzyme cofactor; NAD(P)H; glucose transport; meiosis;
 KW heterokaryon; taxol; cyclosporin A; erythromycin; viral infection;
 KW plant genome shuffling; micropore manipulation; reiterative pooling.

XX OS Escherichia coli.

XX PN AU2004200501-A1.

XX PD 04-MAR-2004.

XX PF 09-FEB-2004; 2004AU-00200501.

XX PR 09-FEB-2004; 2004AU-00200501.

XX PA (MAXY-) MAXYGEN INC.

XX PI Yuan L, Cox T, Bass S, Castle LA, Patten PA, Ness JE, Tobin M;
 PI Affholter JA, Huismann G, Zhang Y, Krebber CM, Subramanian V;
 PI Minshull J, Stemmer WPC, Del Cardayre S;

XX DR WPI; 2004-507924/49.
 XX N-PSDB; ADQ87815.

XX Evolving cells to acquire a desired property, by forming protoplasts of
 XX different cells, fusing protoplasts to form hybrid protoplasts (HP),
 XX producing regenerated cells, forming additional HP and producing
 XX additional regenerated cells.

XX PS Disclosure; Fig 13; 196pp; English.

XX The invention relates to a method of evolving cells to acquire a desired
 CC property by forming protoplasts of different cells, fusing the
 CC protoplasts to form hybrid protoplasts, incubating the hybrid protoplasts
 CC to produce regenerated cells, repeatedly forming protoplasts from
 CC regenerated cells, fusing protoplasts to form hybrid protoplasts in which
 CC genomes from protoplasts recombine to form additional hybrid genomes and
 CC incubating additional hybrid protoplasts for producing additional
 CC regenerated cells. The invention also relates to a method of producing a
 CC library of diverse multicellular organisms involving providing a pool of
 CC male gametes and a pool of female gametes, where one of the male pool or
 CC the female pool comprises several different gametes derived from
 CC different strains of a species or different species and the male gametes

CC fertilise the female gametes, permitting at least a portion of the
CC resulting fertilised gametes to grow into reproductively viable
CC organisms, repeatedly crossing the reproductively viable organisms to
CC produce a library of diverse organisms and selecting the library for a
CC desired trait or property. The methods are useful for evolving cells to
CC acquire desired properties such as heat tolerance, ethanol production,
CC ethanol tolerance, improved production and maintenance of enzyme
CC cofactors, improved production and maintenance of NAD(P)H and improved
CC glucose transport. The desired property is the expression of a protein, or
CC primary metabolite or secondary metabolite, the secretion of a protein or
CC secondary metabolite, capacity for meiosis or compatibility to form a
CC heterokaryon with another strain. The secondary metabolite is chosen from
CC taxol, cyclosporin A and erythromycin. The method is also useful for
CC acquisition of desired properties such as enhanced recombinationicity,
CC gene copy number or gene reductivity and capacity for expression and/or
CC secretion of proteins or secondary metabolites. The method is further
CC useful for predicting efficacy of a drug in treating viral infection,
CC plant genome shuffling, micropore manipulation, producing transgenic
CC animals, improvement of overexpressed genes for a desired phenotype and
CC reiterative pooling and breeding of higher organisms. The cells evolved
CC by the method are useful in molecular genetics. This sequence represents
CC an E. coli recA protein clone used in the method of the invention.
XX
XX Sequence 358 AA;

Query Match 100.0%; Score 1678; DB 8; Length 358;
Best Local Similarity 100.0%; Pred. No. 6.3e-155;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSDMDVETISTGSLDLALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSDMDVETISTGSLDLALGAGGLPMGR 66
QY 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDDIDNLLCSQPD 120
DB 67 IVEIYGPSSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGETGDSHMGIAARMWSQAMRKLKAG 180
DB 127 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGETGDSHMGIAARMWSQAMRKLKAG 186
QY 181 NLKQSTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRIGAVGEGENVVGS 240
DB 187 NLKQSTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRIGAVGEGENVVGS 246
QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLIKAGAWYSYKGEKIQ 300
DB 247 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLIKAGAWYSYKGEKIQ 306
QY 301 GKANATAWLKNDPETAKEIEKKVRELLSNPNSTP 335
DB 307 GKANATAWLKNDPETAKEIEKKVRELLSNPNSTP 341

RESULT 8
ADQ87821
ID ADQ87821 standard; protein; 358 AA.
AC ADQ87821;
XX
XX
XX 09-SEP-2004 (first entry)
XX
XX E. coli recA protein.
XX
XX RecA; protoplast formation; gamete; heat tolerance; ethanol production;
KW ethanol tolerance; enzyme cofactor; NAD(P)H; glucose transport; meiosis;
KW heterokaryon; taxol; cyclosporin A; erythromycin; viral infection;
KW plant genome shuffling; micropore manipulation; reiterative pooling.
XX Escherichia coli.
OS
XX
PN AU2004200501-A1.
XX

PD 04-MAR-2004.
XX
XX 09-FEB-2004; 2004AU-00200501.
XX
XX 09-FEB-2004; 2004AU-00200501.
XX
XX (MAXY-) MAXYGEN INC.
XX
XX Yuan L, Cox T, Bass S, Castle LA, Patten PA, Ness JE, Tobin M;
PI Affholter JA, Huismann G, Zhang Y, Kriebber CW, Subramanian V;
PI Minshull J, Stemmer WPC, Del Cardayre S;
XX
XX WPI: 2004-507924/49.
DR N-PADB; ADQ87814.
XX
XX Evolving cells to acquire a desired property, by forming protoplasts of
PT different cells, fusing protoplasts to form hybrid protoplasts (HP),
PT producing regenerated cells, forming additional HP and producing
PT additional regenerated cells.
XX
XX Disclosure; Fig 13; 196pp; English.
XX
XX The invention relates to a method of evolving cells to acquire a desired
CC property by forming protoplasts of different cells, fusing the
CC protoplasts to form hybrid protoplasts, incubating the hybrid protoplasts
CC to produce regenerated cells, repeatedly forming protoplasts from
CC regenerated cells, fusing protoplasts to form hybrid protoplasts in which
CC genomes from protoplasts recombine to form additional hybrid genomes and
CC incubating additional hybrid protoplasts for producing additional
CC regenerated cells. The invention also relates to a method of producing a
CC library of diverse multicellular organisms involving providing a pool of
CC male gametes and a pool of female gametes, where one of the male pool or
CC the female pool comprises several different gametes derived from
CC different strains of a species or different species and the male gametes
CC fertilise the female gametes, permitting at least a portion of the
CC resulting fertilised gametes to grow into reproductively viable
CC organisms, repeatedly crossing the reproductively viable organisms to
CC produce a library of diverse organisms and selecting the library for a
CC desired trait or property. The methods are useful for evolving cells to
CC acquire desired properties such as heat tolerance, ethanol production,
CC ethanol tolerance, improved production and maintenance of enzyme
CC cofactors, improved production and maintenance of NAD(P)H and improved
CC glucose transport. The desired property is the expression of a protein,
CC primary metabolite or secondary metabolite, the secretion of a protein or
CC secondary metabolite, capacity for meiosis or compatibility to form a
CC heterokaryon with another strain. The secondary metabolite is chosen from
CC taxol, cyclosporin A and erythromycin. The method is also useful for
CC acquisition of desired properties such as enhanced recombinationicity,
CC gene copy number or gene reductivity and capacity for expression and/or
CC secretion of proteins or secondary metabolites. The method is further
CC useful for predicting efficacy of a drug in treating viral infection,
CC plant genome shuffling, micropore manipulation, producing transgenic
CC animals, improvement of overexpressed genes for a desired phenotype and
CC reiterative pooling and breeding of higher organisms. The cells evolved
CC by the method are useful in molecular genetics. This sequence represents
CC the E. coli recA protein used in the method of the invention.
XX
XX Sequence 358 AA;
SQ
Query Match 100.0%; Score 1678; DB 8; Length 358;
Best Local Similarity 100.0%; Pred. No. 6.3e-155;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSDMDVETISTGSLDLALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSDMDVETISTGSLDLALGAGGLPMGR 66
QY 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDDIDNLLCSQPD 120
DB 67 IVEIYGPSSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGETGDSHMGIAARMWSQAMRKLKAG 180
XX

Db 127 TGEQALEICDALSARGAVDVIVVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMKRLAG 186
 QY 181 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVGS 240
 Db 187 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVGS 246
 QY 241 ETRVKVVKNKIAAPFQKAEFQILYGGINFGYBELVGLVKEKLEIKAGAWSYKGEKIQG 300
 Db 247 ETRVKVVKNKIAAPFQKAEFQILYGGINFGYBELVGLVKEKLEIKAGAWSYKGEKIQG 306
 QY 301 GKANATAWLKNDPETAKEIEKKVRELLLSNPSTP 335
 Db 307 GKANATAWLKNDPETAKEIEKKVRELLLSNPSTP 341

RESULT 9

AAV68831
 ID AAV68831 standard; protein; 358 AA.

AC AAV68831;

DT 16-MAY-2000 (first entry)

DE Amino acid sequence of a hyperrecombinogenic recA protein clone 6.

XX recA; hyperrecombinogenic variant; male gamete; female gamete;
 KW heat tolerance; ethanol production; ethanol tolerance; metabolite; taxol;
 KW cyclosporin A; erythromycin; meiosis.

OS Synthetic.
 OS Escherichia coli.

XX Key Location/Qualifiers
 FH Misc-difference 156
 FT /note= "encoded by TCG"

XX WO200004190-A1.
 XX 27-JAN-2000.

PD 15-JUL-1999; 99WO-US015972.

PF 15-JUL-1999; 98US-00116180.

PR (MAXY-) MAXYGEN INC.

XX Del Cardayre S, Tobin M, Stemmer WPC, Ness JB, Minshull J;
 PI Patten PA, Subramanian V, Castle LA, Krebber CM, Bass S, Zhang Y;
 PI Cox T, Huisman G, Yuan L, Alfholter JA;

XX WPI; 2000-182446/16.
 XX N-PSDB; AA260613.

XX Evolution of whole cells and organism by iterative cycles of
 PT recombination and selection and screening for acquisition of desired
 PT properties.

XX Example 1; Fig 13; 197pp; English.

XX The present sequence represents a hyperrecombinogenic recA protein. The
 CC wild type recA protein (AAV68827) was used to produce hyperrecombinogenic
 CC variants, using the method of the invention. The specification describes
 CC a method for producing a library of diverse multicellular organisms using
 CC pools of male and female gametes. At least one of the male pool or female
 CC pools comprises a number of different gametes derived from different
 CC strains of a species or of a different species. The viable organisms
 CC produced from the fertilized gametes are repeatedly crossed to produce a
 CC library of diverse organisms, which are selected for a desired trait or
 CC property. The methods can be used to evolve cells to acquire a desired
 CC property such as heat tolerance, ethanol production or tolerance, acid,
 CC improved production and maintenance of enzyme cofactors or NAD(P)H and
 CC improved glucose transport. The desired property may be expression of a
 CC protein or primary or secondary metabolite. Alternatively the desired

CC property is secretion of a protein or secondary metabolite, chosen from
 CC taxol, cyclosporin A and erythromycin. The desired property may be a
 CC capacity for meiosis or compatibility to form a heterokaryon with another
 CC strain
 XX Sequence 358 AA;

Query Match 99.9%; Score 1677; DB 3; Length 358;
 Best Local Similarity 99.7%; Pred. No. 7.9e-155;
 Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDMDVETISTGSLSLDIALGAGGLPMGR 60
 Db 7 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDMDVETISTGSLSLDIALGAGGLPMGR 66

QY 61 IVEIYGPSSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGYDIDNLLCSQPD 120

Db 67 IVEIYGPSSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGYDIDNLLCSQPD 126

QY 121 TGEQALEICDALSARGAVDVIVVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMKRLAG 180

Db 127 TGEQALEICDALSARGAVDVIVVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMKRLAG 186

QY 181 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVGS 240

Db 187 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVGS 246

QY 241 ETRVKVVKNKIAAPFQKAEFQILYGGINFGYBELVGLVKEKLEIKAGAWSYKGEKIQG 300

Db 247 ETRVKVVKNKIAAPFQKAEFQILYGGINFGYBELVGLVKEKLEIKAGAWSYKGEKIQG 306

QY 301 GKANATAWLKNDPETAKEIEKKVRELLLSNPSTP 335

Db 307 GKANATAWLKNDPETAKEIEKKVRELLLSNPSTP 341

RESULT 10

ADRI6430

ID ADRI6430 standard; protein; 335 AA.

AC ADRI6430;

XX 04-NOV-2004 (first entry)

DE E. coli RecA mutant protein #2.

XX RecA; DNA binding protein; ssDNA binding protein; SSB;
 KW DNA strand exchange reaction; site-specific cleavage; mutant; muten.
 XX Escherichia coli.

XX Key Location/Qualifiers

FT Misc-difference 38

FT /note= "Wild-type Glu substituted by Lys"

PN US2004157248-A1.

XX 12-AUG-2004.

PF 11-DEC-2003; 2003US-00733782.

PR 12-DEC-2002; 2002US-0432758P.

XX (COXM/) COX M M.

PA (LUSE/) LUSSETTI S L.

PA (EGGL/) EGGLER A L.

XX (HARU/) HARUTA N.

PI Cox MM, Lusetti SL, Eggler AL, Haruta N;

XX WPI; 2004-580265/56.

DR N-PSDB; ADRI6431.

PT New RecA mutant proteins comprising a single mutation or a double
PT mutation, useful for catalyzing homologous DNA pairing and DNA strand
PT exchange reactions in an in vitro or in vivo environment.
XX
XX
PS Claim 13; SEQ ID NO 3; 34pp; English.
XX
CC The invention relates to an isolated RecA mutant protein, where the
CC protein is either a single mutant RecA protein comprising a deletion of
CC amino acid residues from the carboxyl terminus, or a double mutant RecA
CC protein comprising a deletion of amino acid residues from the carboxyl
CC terminus and an amino acid change from a glutamate to a basic amino acid.
CC The invention also relates to a polynucleotide sequence that encodes the
CC RecA mutant protein, a method of catalysing an in vitro homologous DNA
CC pairing and DNA strand exchange reactions comprising providing an amount
CC of the RecA mutant protein and a method of increasing recombination
CC efficiency of homologous DNA pairing and DNA strand exchange reactions in
CC a cell comprising supplying to the cell an amount of the RecA mutant
CC protein. The RecA mutant protein comprises an enhanced capacity to
CC displace a DNA binding protein as compared to wild-type RecA. The DNA
CC binding protein is the ssDNA binding protein from Escherichia coli (SSB).
CC It also comprises enhanced binding to DNA during a DNA strand exchange
CC reaction as compared to wild-type RecA. The composition and methods are
CC useful for catalyzing homologous DNA pairing and DNA strand exchange
CC reactions in an in vitro or in vivo environment. These may be used in
CC promoting in vitro alterations of genes to permit the rapid construction
CC of desired gene mutants for industrial and pharmaceutical purposes. The
CC mutant proteins may also be used in targeting the site-specific cleavage
CC of small and large DNAs, or as a basis for the design and construction of
CC tiny electronic circuits based on DNA. This sequence represents a RecA
CC mutant protein of the invention.
XX
SQ Sequence 335 AA;

Query Match 99.8%; Score 1674; DB 8; Length 335;
Best Local Similarity 99.7%; Pred. No. 1.4e-154;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLDLIALGAGGLPMGR 60
DB 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLDLIALGAGGLPMGR 60
QY 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCFIDEHALDPIYARKLGVDIDNLLCSQPD 120
DB 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCFIDEHALDPIYARKLGVDIDNLLCSQPD 120
QY 121 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGEIGDSHMGLAARMWSQAMKRLAG 180
DB 121 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGEIGDSHMGLAARMWSQAMKRLAG 180
QY 181 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
DB 181 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKKEKLEKAGAWSYKGEKIGQ 300
DB 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKKEKLEKAGAWSYKGEKIGQ 300
QY 301 GKANATAWLKDNPETAKEIEKKVRELLLSNPSTP 335
DB 301 GKANATAWLKDNPETAKEIEKKVRELLLSNPSTP 335

RESULT 11
AAW64213
ID AAW64213 standard; protein; 358 AA.
XX
AC AAW64213;
XX

28-APR-1999 (first entry)

New minshall recA protein.

Escherichia coli; recA protein; hyperrecombinogenic variant; Minshall;

KW recursive sequence recombination; evolution.
XX Escherichia coli.
XX WO9831837-A1.
XX
XX 23-JUL-1998.
XX
XX 16-JAN-1998; 98WO-US0000852.
XX
XX 17-JAN-1997; 97US-0035054P.
XX
XX (MAXY-) MAXYGEN INC.
XX
XX Delcardayre SB, Tobin MB, Stemmer WPC, Ness JE, Minshall J;
PI Patten P;
XX
XX WPI; 1998-427565/60.
XX N-PSDB; AAV44285.
XX
XX Evolution of whole cells and organisms by recursive DNA sequence
XX recombination in cells to evolve cells having acquired desired function,
XX useful in methods for predicting the efficacy of a drug in treating viral
XX or pathogenic infections.
XX
XX Example 1; Fig 13; 125pp; English.
XX
XX The present invention provides methods employing iterative cycles of
XX recombination and selection/screening for evolution of whole cells and
XX organisms toward acquisition of desired properties, e.g. enhanced
XX recombinogenicity, genome copy number, and capacity for expression and/or
XX secretion of proteins and secondary metabolites. The present sequence
XX represents a wild-type recA protein (designated new Minshall), from an
XX example of the present invention
XX
SQ Sequence 358 AA;

Query Match 99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 1.6e-154;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLDLIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLDLIALGAGGLPMGR 66
QY 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCFIDEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPSSGKTTLTQVIAAAQREGKTCFIDEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGEIGDSHMGLAARMWSQAMKRLAG 180
DB 127 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGEIGDSHMGLAARMWSQAMKRLAG 186
QY 181 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
DB 187 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 246
QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKKEKLEKAGAWSYKGEKIGQ 300
DB 247 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKKEKLEKAGAWSYKGEKIGQ 306
QY 301 GKANATAWLKDNPETAKEIEKKVRELLLSNPSTP 335
DB 307 GKANATAWLKDNPETAKEIEKKVRELLLSNPSTP 341

RESULT 12
AAV68828
ID AAV68828 standard; protein; 358 AA.
XX
XX AAV68828;
XX
XX 16-MAY-2000 (first entry)
DT

XX Amino acid sequence of a hyperrecombinogenic recA protein clone 2.
DE recA; hyperrecombinogenic variant; male gamete; female gamete;
KW heat tolerance; ethanol production; ethanol tolerance; metabolite; taxol;
KW cyclosporin A; erythromycin; meiosis.
XX Synthetic.
OS Escherichia coli.
XX WO200004190-A1.
XX 27-JAN-2000.
PD 15-JUL-1999; 99WO-US015972.
PF 15-JUL-1998; 98US-00116188.
PR 15-JUL-1998; 98US-00116188.
XX (MAXY-) MAXYGEN INC.
XX Del Cardayre S, Tobin M, Stemmer WPC, Ness JE, Minshull J; Zhang Y;
PI Patten PA, Subramanian V, Castle LA, Krebber CM, Bass S, Zhang Y;
PI Cox T, Huisman G, Yuan L, Affholter JA;
XX WPI; 2000-182446/16.
DR N-PSDB; AAZ60610.
XX Evolution of whole cells and organism by iterative cycles of
PT recombination and selection and screening for acquisition of desired
PT properties.
XX Example 1; Fig 13; 197pp; English.
XX The present sequence represents a hyperrecombinogenic recA protein. The
CC wild type recA protein (AAV68827) was used to produce hyperrecombinogenic
CC variants, using the method of the invention. The specification describes
CC a method for producing a library of diverse multicellular organisms using
CC pools of male and female gametes. At least one of the male pool or female
CC pools comprises a number of different gametes derived from different
CC strains of a species or of a different species. The viable organisms
CC produced from the fertilized gametes are repeatedly crossed to produce a
CC library of diverse organisms, which are selected for a desired trait or
CC property. The methods can be used to evolve cells to acquire a desired
CC property such as heat tolerance, ethanol production or tolerance, acid,
CC improved production and maintenance of enzyme cofactors or NAD(P)H and
CC improved glucose transport. The desired property may be expression of a
CC protein or primary or secondary metabolite. Alternatively the desired
CC property is secretion of a protein or secondary metabolite, chosen from
CC taxol, cyclosporin A and erythromycin. The desired property may be a
CC capacity for meiosis or compatibility to form a heterokaryon with another
CC strain
XX SQ Sequence 358 AA;
Query Match 99.7%; Score 1673; DB 3; Length 358;
Best Local Similarity 99.4%; Pred. No. 1.9e-154;
Matches 333; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 AIDENKQALAAALGQIEKQFGKSGIMRLGEDRSMVETITSGSLDLALGAGLPMGR 60
DB 7 AIDENKQALATGALGQIEKQFGKSGIMRLGEDRSMVETITSGSLDLALGAGLPMGR 66
QY 61 IVEIYGPESSGKTTLTLOVIAAAREGKTCFIDAEHALDPIYARKLGVDTNLLCSQPD 120
DB 67 IVEIYGPESSGKTTLTLOVIAAAREGKTCFIDAEHALDPIYARKLGVDTNLLCSQPD 126
QY 121 TGEQALEICDALARGAVDVIWDSVAALTTPKAEIEGEGISGHMGLAARMSQAMRKLKAG 180
DB 127 TGEQALEICDALARGAVDVIWDSVAALTTPKAEIEGEGISGHMGLAARMSQAMRKLKAG 186
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNAKLFYASVRLDIRRIGAVKEGENVVGVS 240
DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNAKLFYASVRLDIRRIGAVKEGENVVGVS 246

QY 241 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKELIEKAGAWSYKGEKIQG 300
DB 247 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKELIEKAGAWSYKGEKIQG 306
QY 301 GKANATAWLKONPETAKIEKKVRELLSNPNSTP 335
DB 307 GKANATAWLKONPETAKIEKKVRELLSNPNSTP 341
RESULT 13
ADQ87823
ID ADQ87823 standard; protein; 358 AA.
XX AC ADQ87823;
XX 09-SEP-2004 (first entry)
XX E. coli recA protein clone #2.
XX RecA; protoplast formation; gamete; heat tolerance; ethanol production;
KW ethanol tolerance; enzyme cofactor; NAD(P)H; glucose transport; meiosis;
KW heterokaryon; taxol; cyclosporin A; erythromycin; viral infection;
KW plant genome shuffling; micropore manipulation; reiterative pooling.
XX OS Escherichia coli.
XX FN AU2004200501-A1.
XX PD 04-MAR-2004.
XX 09-FEB-2004; 2004AU-00200501.
XX 09-FEB-2004; 2004AU-00200501.
XX (MAXY-) MAXYGEN INC.
XX Yuan L, Cox T, Bass S, Castle LA, Patten PA, Ness JE, Tobin M;
PI Affholter JA, Huisman G, Zhang Y, Krebber CM, Subramanian V;
PI Minshull J, Stemmer WPC, Del Cardayre S;
XX WPI; 2004-507924/49.
DR N-PSDB; ADQ87816.
XX Evolving cells to acquire a desired property, by forming protoplasts of
PT different cells, fusing protoplasts to form hybrid protoplasts (HP),
PT producing regenerated cells, forming additional HP and producing
PT additional regenerated cells.
XX Disclosure; Fig 13; 196pp; English.
XX The invention relates to a method of evolving cells to acquire a desired
CC property by forming protoplasts of different cells, fusing the
CC protoplasts to form hybrid protoplasts, incubating the hybrid protoplasts
CC to produce regenerated cells, repeatedly forming protoplasts from
CC regenerated cells, fusing protoplasts to form hybrid protoplasts in which
CC genomes from protoplasts recombine to form additional hybrid genomes and
CC incubating additional hybrid protoplasts for producing additional
CC regenerated cells. The invention also relates to a method of producing a
CC library of diverse multicellular organisms involving providing a pool of
CC male gametes and a pool of female gametes, where one of the male pool or
CC the female pool comprises several different gametes derived from
CC different strains of a species or different species and the male gametes
CC fertilise the female gametes, permitting at least a portion of the
CC resulting fertilised gametes to grow into reproductively viable
CC organisms, repeatedly crossing the reproductively viable organisms to
CC produce a library of diverse organisms and selecting the library for a
CC desired trait or property. The methods are useful for evolving cells to
CC acquire desired properties such as heat tolerance, ethanol production,
CC ethanol tolerance, improved production and maintenance of enzyme
CC cofactors, improved production and maintenance of NAD(P)H and improved
CC glucose transport. The desired property is the expression of a protein, or
CC primary metabolite or secondary metabolite, the secretion of a protein or

CC secondary metabolite, capacity for meiosis or compatibility to form a
CC heterokaryon with another strain. The secondary metabolite is chosen from
CC taxol, cyclosporin A and erythromycin. The method is also useful for
CC acquisition of desired properties such as enhanced recombination, or
CC gene copy number or gene reductivity and capacity for expression and/or
CC secretion of proteins or secondary metabolites. The method is further
CC useful for predicting efficacy of a drug in treating viral infection,
CC plant genome shuffling, micropore manipulation, producing transgenic
CC animals, improvement of overexpressed genes for a desired phenotype and
CC reiterative pooling and breeding of higher organisms. The cells evolved
CC by the method are useful in molecular genetics. This sequence represents
CC an E. coli recA protein clone used in the method of the invention.

XX Sequence 358 AA;

Query Match 99.7%; Score 1673; DB 8; Length 358;
Best Local Similarity 99.4%; Pred. No. 1.9e-154;
Matches 333; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLDLALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLDLALGAGGLPMGR 66
QY 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDDIDNLLCSQPD 120
DB 67 IVEIYGPSSGKTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEIGDSHGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEIGDSHGLAARMMSQAMRKLKAG 186
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPDPTTGGNALKFYASVRLDIRRIGAVGEGENVVGS 240
DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPDPTTGGNALKFYASVRLDIRRIGAVGEGENVVGS 246
QY 241 ETRKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLIIEKAGAWSYKGEKIQG 300
DB 247 ETRKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLIIEKAGAWSYKGEKIQG 306
QY 301 GRANATAWLKNDPETAKEKKVRELLSNPNSTP 335
DB 307 GRANATAWLKNDPETAKEKKVRELLSNPNSTP 341

RESULT 14
ADQ87826
ID ADQ87826 standard; protein; 358 AA.

XX ADQ87826;
XX
XX 09-SEP-2004 (first entry)
XX
DE E. coli recA protein clone #5.
XX
XX RecA; protoplast formation; gamete; heat tolerance; ethanol production;
XX ethanol tolerance; enzyme cofactor; NAD(P)H; glucose transport; meiosis;
KW heterokaryon; taxol; cyclosporin A; erythromycin; viral infection;
KW plant genome shuffling; micropore manipulation; reiterative pooling.
XX Escherichia coli.
OS
XX AU2004200501-A1.
XX
XX 04-MAR-2004.
XX
XX 09-FEB-2004; 2004AU-00200501.
XX
XX 09-FEB-2004; 2004AU-00200501.
XX (MAXY-) MAXYGEN INC.
XX Yuan L, Cox T, Bass S, Castle LA, Patten PA, Ness JE, Tobin M;
PI Affholter JA, Huisman G, Zhang Y, Krebber CM, Subramanian V;

PI

Minshull J, Stemmer WPC, Del Cardayre S;

WPI; 2004-507924/49.
N-PSDB; ADQ87819.

XX Evolving cells to acquire a desired property, by forming protoplasts of
XX different cells, fusing protoplasts to form hybrid protoplasts (HP),
XX producing regenerated cells, forming additional HP and producing
XX additional regenerated cells.

PS Disclosure; Fig 13; 196pp; English.

XX The invention relates to a method of evolving cells to acquire a desired
XX property by forming protoplasts of different cells, fusing the
XX protoplasts to form hybrid protoplasts, incubating the hybrid protoplasts
XX to produce regenerated cells, repeatedly forming protoplasts from
XX regenerated cells, fusing protoplasts to form hybrid protoplasts in which
XX genomes from protoplasts recombine to form additional hybrid genomes and
XX incubating additional hybrid protoplasts for producing additional
XX regenerated cells. The invention also relates to a method of producing a
XX library of diverse multicellular organisms involving providing a pool of
XX male gametes and a pool of female gametes, where one of the male pool or
XX the female pool comprises several different gametes derived from
XX different strains of a species or different species and the male gametes
XX fertilise the female gametes, permitting at least a portion of the
XX resulting fertilised gametes to grow into reproductively viable
XX organisms, repeatedly crossing the reproductively viable organisms to
XX produce a library of diverse organisms and selecting the library for a
XX desired trait or property. The methods are useful for evolving cells to
XX acquire desired properties such as heat tolerance, ethanol production,
XX ethanol tolerance, improved production and maintenance of enzyme
XX cofactors, improved production and maintenance of NAD(P)H and improved
XX glucose transport. The desired property is the expression of a protein,
XX primary metabolite or secondary metabolite, the secretion of a protein or
XX heterokaryon with another strain. The secondary metabolite is chosen from
XX taxol, cyclosporin A and erythromycin. The method is also useful for
XX acquisition of desired properties such as enhanced recombination, or
XX gene copy number or gene reductivity and capacity for expression and/or
XX secretion of proteins or secondary metabolites. The method is further
XX useful for predicting efficacy of a drug in treating viral infection,
XX plant genome shuffling, micropore manipulation, producing transgenic
XX animals, improvement of overexpressed genes for a desired phenotype and
XX reiterative pooling and breeding of higher organisms. The cells evolved
XX by the method are useful in molecular genetics. This sequence represents
XX an E. coli recA protein clone used in the method of the invention.

XX Sequence 358 AA;

Query Match 99.5%; Score 1669; DB 8; Length 358;
Best Local Similarity 99.4%; Pred. No. 4.8e-154;
Matches 333; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLDLALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLDLALGAGGLPMGR 66
QY 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDDIDNLLCSQPD 120
DB 67 IVEIYGPSSGKTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEIGDSHGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEIGDSHGLAARMMSQAMRKLKAG 186
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPDPTTGGNALKFYASVRLDIRRIGAVGEGENVVGS 240
DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPDPTTGGNALKFYASVRLDIRRIGAVGEGENVVGS 246
QY 241 ETRKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLIIEKAGAWSYKGEKIQG 300
DB 247 ETRKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLIIEKAGAWSYKGEKIQG 306

QY 301 GKANATAWLKONPETAKIEKKVRELLLSNPNSTP 335
DB 307 GKANATAWLKONPETAKIEKKVRELLLSNPNSTP 341

RESULT 15

AAW64215
ID AAW64215 standard; protein; 358 AA.

XX AAW64215;
XX AC
XX 28-APR-1999 (first entry)
XX DE Hyperrecombinogenic variant recA protein clone 4.
XX KW Escherichia coli; recA protein; hyperrecombinogenic variant; Minshall;
XX OS recursive sequence recombination; evolution.
XX OS Escherichia coli.
XX OS Synthetic.
XX PN WO9831837-A1.
XX PD 23-JUL-1998.
XX PF 16-JAN-1998; 98WO-US000852.
XX PR 17-JAN-1997; 97US-0035054P.
XX PA (MAXY-) MAXYGEN INC.
XX PI Delcardayre SB, Tobin MB, Stemmer WPC, Ness JE, Minshall J;
XX PI Patten P;
XX DR WPI: 1998-427565/60.
XX DR N-PSDB; AAV44287.

PT Evolution of whole cells and organisms by recursive DNA sequence
PT recombination in cells to evolve cells having acquired desired function,
PT useful in methods for predicting the efficacy of a drug in treating viral
PT or pathogenic infections.

PS Example 1; Fig 13; 125pp; English.

XX The present invention provides methods employing iterative cycles of
CC recombination and selection/screening for evolution of whole cells and
CC organisms toward acquisition of desired properties, e.g. enhanced
CC recombination, genome copy number, and capacity for expression and/or
CC secretion of proteins and secondary metabolites. The present sequence
CC represents a hyperrecombinogenic variant of a recA protein (see
CC AAW64213), from an example of the present invention

XX SQ Sequence 358 AA;

Query Match 99.4%; Score 1668; DB 2; Length 358;
Best Local Similarity 99.1%; Pred. No. 6e-154;
Matches 332; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 66
QY 61 IVEIYGPSGGKTTLTQVIAAQRGKTCAPIDAEHALDPIYARKLGVDIDNLCSQPD 120
DB 67 IVEIYGPSGGKTTLTQVIAAQRGKTCAPIDAEHALDPIYARKLGVDIDNLCSQPD 126
QY 121 TGEQALEICDALARGSGVDVIVDSVAALTPKAEIEGIGDGHMGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARGSGVDVIVDSVAALTPKAEIEGIGDGHMGLAARMMSQAMRKLKAG 186
QY 181 NLKQNTLLIFINQIRMKIGVFNPGPPTTGGNALKFYASVRLDIRRIGAVKEGENVYGS 240
DB 187 NLKQNTLLIFINQIRMKIGVFNPGPPTTGGNALKFYASVRLDIRRIGAVKEGENVYGS 246

QY 241 ETRVKVVKNNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLIBKAGAWYSYKGEKIGQ 300
DB 247 ETRVKVVKNNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLIBKAGAWYSYKGEKIGQ 306
QY 301 GKANATAWLKONPETAKIEKKVRELLLSNPNSTP 335
DB 307 GKANATAWLKONPETAKIEKKVRELLLSNPNSTP 341

Search completed: February 16, 2006, 01:18:51
Job time : 135 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 01:19:09 ; Search time 25.5 Seconds
(without alignments)
1264.024 Million cell updates/sec

Title: US-10-733-782-1
Perfect score: 1678
Sequence: 1 AIDENKOKALAAALGQIEKQ.....AKETIKKVELLSNPSTP 335
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1678	100.0	353	1 ROECA	recombination prot
2	1678	100.0	353	2 AH2806	RecA protein [impo
3	1666	99.3	353	2 H85917	hypothetical prote
4	1641	97.8	353	2 AG0843	RecA protein [impo
5	1534	91.4	356	2 AG0401	RecA protein [impo
6	1532	91.3	354	2 S31481	recombination prot
7	1530	91.2	356	2 S37586	recombination prot
8	1504	89.6	355	1 ROEBPM	recombination prot
9	1436	85.6	412	2 E82310	recA protein VC054
10	1417	84.4	354	2 S46274	recombination prot
11	1412.5	84.2	354	2 JQ1461	recombination prot
12	1332	79.4	358	2 A49929	recombination prot
13	1323	78.8	353	2 JC5198	recombination prot
14	1268	75.6	346	1 RQPSAA	RecA protein PA361
15	1265	75.4	349	2 JN0321	recombination prot
16	1264.5	75.4	344	2 JQ0776	recombination prot
17	1257	74.9	349	2 I39509	recombination prot
18	1255	74.8	355	2 T10482	recA protein - Pse
19	1250	74.5	352	2 JC2406	recombination prot
20	1246	74.3	352	2 A47709	recombination prot
21	1238.5	73.8	348	1 A60989	recombination prot
22	1237.5	73.2	342	2 S18728	recombination prot
23	1217.5	72.6	348	2 F81082	RecA protein NMB14
24	1216.5	72.5	348	2 F81860	RecA protein NMA16
25	1213	72.3	378	2 AE3350	recA protein [impo
26	1203.5	71.7	348	2 JQ0745	recombination prot
27	1203.5	71.7	351	2 S16896	recombination prot
28	1201.5	71.6	347	1 RQPSAC	recombination prot
29	1200.5	71.5	363	2 JC1377	recombination prot

ALIGNMENTS

RESULT 1

ROECA

recombination protein recA [validated] - Escherichia coli (strain K-12)

N:Alternate names: recombinase A

C:Species: Escherichia coli

C>Date: 31-Jul-1980 #sequence revision 14-Nov-1997 #text change 09-Jul-2004

C:Accession: G65049; A93847; A93846; S11931; S63525; S69129; S63979; A03548

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: G65049

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-353 <BLAT>

A:Cross-references: UNIPROT:P03017; UNIPARC:UPI0000112C1C; GB:AE000354; GB:U00096; NID

A:Experimental source: strain K-12, substrain MG1655

R:Horii, T.; Ogawa, T.; Ogawa, H.

Proc. Natl. Acad. Sci. U.S.A. 77, 313-317, 1980

A:Title: Organization of the recA gene of Escherichia coli.

A:Reference number: A93847; MUID:80145618; PMID:6244554

A:Accession: A93847

A:Molecule type: DNA

A:Residues: 2-353 <HOR>

A:Cross-references: UNIPARC:UPI0000112C17; GB:V00328; NID:942672; PIDN:CAA23618.1; PID

R:Sancar, A.; Stachelek, C.; Konigsberg, W.; Rupp, W.D.

Proc. Natl. Acad. Sci. U.S.A. 77, 2611-2615, 1980

A:Title: Sequences of the recA gene and protein.

A:Reference number: A93846; MUID:80234673; PMID:6930655

A:Accession: A93846

A:Molecule type: DNA

A:Residues: 2-353 <SAN>

A:Cross-references: UNIPARC:UPI0000112C17; GB:V00328; NID:942672; PIDN:CAA23618.1; PIDN:CAA23618.1

R:Zhao, X.J.; McEntee, K.

Mol. Gen. Genet. 222, 369-376, 1990

A:Title: DNA sequence analysis of the recA genes from Proteus vulgaris, Erwinia carotov

A:Reference number: S11931; MUID:91109725; PMID:2274037

A:Accession: S11931

A:Molecule type: DNA

A:Residues: 2-353 <ZHA>

A:Cross-references: UNIPARC:UPI0000112C17; GB:X55552; NID:942678

A>Note: this ORF is not annotated in GenBank entry ECRECAGEN, release 109.0

R:Morimatsu, K.; Horii, T.

Eur. J. Biochem. 234, 695-705, 1995

A:Title: DNA-binding surface of RecA protein. Photochemical cross-linking of the first

A:Reference number: S63525; MUID:96096753; PMID:8529655

A:Accession: S63525

A>Status: preliminary

A:Molecule type: protein

A:Residues: 65-69;90-97;179-184;200-207;258-265;304-311;323-331 <MORI>

A:Cross-references: UNIPARC:UPI00001748A7; UNIPARC:UPI00001748A8; UNIPARC:UPI00001748A9

R;Morimatsu, K.; Horii, T.
Eur. J. Biochem. 228, 772-778, 1995
A:Title: The DNA-binding site of the RecA protein. Photochemical cross-linking of Tyr103
A:Reference number: S69129; MUID:95255284; PMID:7737176
A:Accession: D91073
A:Status: preliminary
A:Molecule type: protein
A:Residues: 90-108,180-184 <MOR2>
A:Cross-references: UNIPARC:UPI00001748AE
R;Gardner, R.V.; Voloshin, O.N.; Camerini-Otero, R.D.
Eur. J. Biochem. 233, 419-425, 1995
A:Title: The identification of the single-stranded DNA-binding domain of the Escherichia
A:Reference number: S63979; MUID:96067680; PMID:7588783
A:Accession: S63979
A:Molecule type: protein
A:Residues: 'XX',187-190,192-194 <GAR>
A:Cross-references: UNIPARC:UPI00001748AF
R;Yu, X.; Egelman, E.H.
submitted to the Brookhaven Protein Data Bank, December 1996
A:Reference number: A67277; PDB:2REC
A:Contents: annotation; X-ray crystallography, 2.3 angstroms, residues 4-157;166-195;211
R;Aihara, H.; Ito, Y.; Kurumizaka, H.; Terada, T.; Yokoyama, S.; Shibata, T.
submitted to the Brookhaven Protein Data Bank, January 1997
A:Reference number: A67455; PDB:1AA3
A:Contents: annotation; conformation by (1)H- and (15)N-NMR, residues 269-331
C:Genetics:
A:Gene: recA
A:Map position: 58 min
C:Function:
A:Description: plays an essential role in homologous recombination, in induction of the
C:Superfamily: recombination protein recA
C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop;
F:167-74/Region: nucleotide-binding motif A (P-loop)
F:141-146/Region: nucleotide-binding motif B
F:73/Binding site: ATP (Lys) #status predicted

Query Match 100.0%; Score 1678; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 4.6e-106; Mismatches 0; Indels 0; Gaps 0;
Matches 335; Conservative 0

Qy 1 AIDENKOKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 61

Qy 61 IVEIYGPSSSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 121

Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGCLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGCLAARMMSQAMRKLKAG 181

Qy 181 NLKQSNLTLLIFINQIRMKIGWFGNPETTTGKNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 182 NLKQSNLTLLIFINQIRMKIGWFGNPETTTGKNALKFYASVRLDIRRIGAVKEGENVVGS 241

Qy 241 ETRVKVKNKIAAPFKQAEFQILYEGEINFGYELVDLGKVEKLEKAGAWTSYKGEKIQG 300
Db 242 ETRVKVKNKIAAPFKQAEFQILYEGEINFGYELVDLGKVEKLEKAGAWTSYKGEKIQG 301

Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 302 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 336

RESULT 2
D91073
RecA protein [imported] - Escherichia coli (strain O157:H7, substrain RMD 0509952)
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: D91073
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: D91073
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <HAY>
A:Cross-references: UNIPROT:P03017; UNIPARC:UPI0000112C1C; GB:BA000007; PIDN:BA036979.1,
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: EC63556
C:Superfamily: recombination protein recA

Query Match 100.0%; Score 1678; DB 2; Length 353;
Best Local Similarity 100.0%; Pred. No. 4.6e-106; Mismatches 0; Indels 0; Gaps 0;
Matches 335; Conservative 0

Qy 1 AIDENKOKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 61

Qy 61 IVEIYGPSSSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 121

Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGCLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGCLAARMMSQAMRKLKAG 181

Qy 181 NLKQSNLTLLIFINQIRMKIGWFGNPETTTGKNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 182 NLKQSNLTLLIFINQIRMKIGWFGNPETTTGKNALKFYASVRLDIRRIGAVKEGENVVGS 241

Qy 241 ETRVKVKNKIAAPFKQAEFQILYEGEINFGYELVDLGKVEKLEKAGAWTSYKGEKIQG 300
Db 242 ETRVKVKNKIAAPFKQAEFQILYEGEINFGYELVDLGKVEKLEKAGAWTSYKGEKIQG 301

Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 302 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 336

RESULT 3
H85917
hypothetical protein recA [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: H85917
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: H85917
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <STO>
A:Cross-references: UNIPARC:UPI00001658E3; GB:AB005174; NID:gl2517139; PIDN:AAG57804.1;
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: recA
C:Superfamily: recombination protein recA

Query Match 99.3%; Score 1666; DB 2; Length 353;
Best Local Similarity 99.4%; Pred. No. 3e-105; Mismatches 1; Indels 0; Gaps 0;
Matches 333; Conservative 1

Qy 1 AIDENKOKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 61

Qy 61 IVEIYGPSSSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120

Db 62 IVEIYGPSSGKTTTLTQVIAAQRGKTCAFIDAEHALDPYARKLGVDIDNLLCSQPD 121
Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIDSHGMLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIDSHGMLAARMMSQAMRKLKAG 181
Qy 181 NLKQSTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240
Db 182 NLKQSTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 241
Qy 241 ETRVKVVKNTKIAAPFKQAEFOLLYGEGINFYGVGLVGLVKEKLEKAGAWTSYKGEKIGQ 300
Db 242 ETRVKVVKNTKIAAPFKQAEFOLLYGEGINFYGVGLVGLVKEKLEKAGAWTSYKGEKIGQ 301
Qy 301 GKANATLWLNKPNPTAKEIEKKVRELLLSNPNSTP 335
Db 302 GKANATLWLNKPNPTAKEIEKKVRELLLSNPNSTP 336

RESULT 4
AH0843
ReCA protein [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain CT18
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AH0843
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001;
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AH0843
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <PAR>
A:Cross-references: UNIPARC:UPI000005A334; GB:AL513382; PIDN:CAD05935.1; PID:g16503906;
C:Gene: STY2950
C:Superfamily: recombination protein recA

Query Match 97.8%; Score 1641; DB 2; Length 353;
Best Local Similarity 97.3%; Pred. No. 1.4e-103;
Matches 326; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGOIEKQFGKGSIMRLGEDRSMDVETISTGSLSDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGOIEKQFGKGSIMRLGEDRSMDVETISTGSLSDIALGAGGLPMGR 61
Qy 61 IVEIYGPSSGKTTTLTQVIAAQRGKTCAFIDAEHALDPYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSGKTTTLTQVIAAQRGKTCAFIDAEHALDPYARKLGVDIDNLLCSQPD 121
Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIDSHGMLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIDSHGMLAARMMSQAMRKLKAG 181
Qy 181 NLKQSTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240
Db 182 NLKQSTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 241
Qy 241 ETRVKVVKNTKIAAPFKQAEFOLLYGEGINFYGVGLVGLVKEKLEKAGAWTSYKGEKIGQ 300
Db 242 ETRVKVVKNTKIAAPFKQAEFOLLYGEGINFYGVGLVGLVKEKLEKAGAWTSYKGEKIGQ 301
Qy 301 GKANATLWLNKPNPTAKEIEKKVRELLLSNPNSTP 335
Db 302 GKANATLWLNKPNPTAKEIEKKVRELLLSNPNSTP 336

RESULT 5
AG0401

RecA protein [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AG0401
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
Li, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whithead, S.; Barrall,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AG0401
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-356 <KUR>
A:Cross-references: UNIPROT:P37858; UNIPARC:UPI0000165AF8; GB:AL590842; PIDN:CAC92539;
C:Gene: recA
C:Superfamily: recombination protein recA

Query Match 91.4%; Score 1534; DB 2; Length 356;
Best Local Similarity 91.9%; Pred. No. 2.4e-96;
Matches 305; Conservative 14; Mismatches 13; Indels 0; Gaps 0;
Qy 1 AIDENKOKALAAALGOIEKQFGKGSIMRLGEDRSMDVETISTGSLSDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGOIEKQFGKGSIMRLGEDRSMDVETISTGSLSDIALGAGGLPMGR 61
Qy 61 IVEIYGPSSGKTTTLTQVIAAQRGKTCAFIDAEHALDPYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSGKTTTLTQVIAAQRGKTCAFIDAEHALDPYARKLGVDIDNLLCSQPD 121
Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIDSHGMLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIDSHGMLAARMMSQAMRKLKAG 181
Qy 181 NLKQSTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240
Db 182 NLKQSTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 241
Qy 241 ETRVKVVKNTKIAAPFKQAEFOLLYGEGINFYGVGLVGLVKEKLEKAGAWTSYKGEKIGQ 300
Db 242 ETRVKVVKNTKIAAPFKQAEFOLLYGEGINFYGVGLVGLVKEKLEKAGAWTSYKGEKIGQ 301
Qy 301 GKANATLWLNKPNPTAKEIEKKVRELLLSNPN 332
Db 302 GKANATLWLNKPNPTAKEIEKKVRELLLSNPN 333

RESULT 6
S31481
recombination protein recA - Enterobacter agglomerans
N:Alternate names: recombinase A
C:Species: Enterobacter agglomerans
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S31481
R:Rappold, C.S.J.; Klingmueller, W.
submitted to the EMBL Data Library, January 1993
A:Description: Cloning and sequencing of the recA gene from Enterobacter agglomerans 33
A:Reference number: S31480
A:Accession: S31481
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-354 <RAP>
A:Cross-references: UNIPROT:P33037; UNIPARC:UPI0000016BC3A; GB:L03291; EMBL:Z19517; NID:
C:Gene: recA
C:Superfamily: recombination protein recA
C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop
P:67-74/Region: nucleotide-binding motif A (P-loop)
P:141-146/Region: nucleotide-binding motif B
F:73/Binding site: ATP (Lys) #status predicted
Query Match 91.3%; Score 1532; DB 2; Length 354;

Best Local Similarity 92.1%; Pred. No. 3.3e-96;
Matches 303; Conservative 15; Mismatches 11; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 61

Qy 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 121

Qy 121 TGEQALEICDALARGAVDVIVDSVAALTTPKAEIEGEIGDSHMGGLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALTRSGAVDVIVDSVAALTTPKAEIEGEIGDSHMGGLAARMMSQAMRKLKAG 181

Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
Db 182 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 241

Qy 241 ETRVKVKNKTAAPKQAEFQILYCEGINFYGELVDLGKVKLEKAGAWYSYKGEKIQG 300
Db 242 ETRVKVKNKTAAPKQAEFQILYCEGINFYGELVDLGKVKLEKAGAWYSYKGEKIQG 301

Qy 301 GKANATAWLKNPETAKEIEKKVRELLLS 329
Db 302 GKANSNVLYKENPKVAELDKKLRDMILLS 330

RESULT 7
S37586
recombination protein recA - Yersinia pestis
N;Alternate names: recombinase A
C;Species: Yersinia pestis
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S37586
submitted to the EMBL Data Library, October 1993
A;Description: Complete nucleotide sequence of Yersinia pestis recA gene.
A;Reference number: S37586
A;Accession: S37586
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-356 <KRV>
A;Cross-references: UNIPROT:P37858; UNIPARC:UPI0000170617; EMBL:X75336; NID:g406793; PID:g406793; PID:g406793

C;Genetics:
A;Gene: recA
C;Superfamily: recombination protein recA
C;Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop;
F;67-74/Region: nucleotide-binding motif A (P-loop)
F;141-146/Region: nucleotide-binding motif B
F;7/3/Binding site: ATP (Lys) #status predicted

Query Match 91.2%; Score 1530; DB 2; Length 356;
Best Local Similarity 91.9%; Pred. No. 4.6e-96;
Matches 305; Conservative 14; Mismatches 13; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 61

Qy 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 121

Qy 121 TGEQALEICDALARGAVDVIVDSVAALTTPKAEIEGEIGDSHMGGLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALTRSGAVDVIVDSVAALTTPKAEIEGEIGDSHMGGLAARMMSQAMRKLKAG 181

Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
Db 182 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 241

Qy 241 ETRVKVKNKTAAPKQAEFQILYCEGINFYGELVDLGKVKLEKAGAWYSYKGEKIQG 300

Db 242 ETRVKVKNKTAAPKQAEFQILYCEGINFYGELVDLGKVKLEKAGAWYSYKGEKIQG 301

Qy 301 GKANATAWLKNPETAKEIEKKVRELLLSNPN 332
Db 302 GKANASVLYKENPANAAELDKKLRMLNGN 333

RESULT 8
ROEBPM
recombination protein recA - Proteus mirabilis
N;Alternate names: recombinase A
C;Species: Proteus mirabilis
C;Date: 30-Sep-1990 #sequence_revision 30-Sep-1990 #text_change 09-Jul-2004
C;Accession: S04606
R;Akaboshi, E.; Yip, M.L.R.; Howard-Flanders, P.
Nucleic Acids Res. 17, 4390, 1989
A;Title: Nucleotide sequence of the recA gene of Proteus mirabilis.
A;Reference number: S04606; MUID:89296502; PMID:2544862
A;Accession: S04606
A;Molecule type: DNA
A;Residues: 1-355 <AKA>
A;Cross-references: UNIPROT:P11406; UNIPARC:UPI000016FD85; GB:X14870; NID:g45631; PIDN:C14870

C;Genetics:
A;Gene: recA
C;Function:
A;Description: plays an essential role in homologous recombination, in induction of the
C;Superfamily: recombination protein recA
C;Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop;
F;67-74/Region: nucleotide-binding motif A (P-loop)
F;141-146/Region: nucleotide-binding motif B
F;7/3/Binding site: ATP (Lys) #status predicted

Query Match 89.6%; Score 1504; DB 1; Length 355;
Best Local Similarity 89.4%; Pred. No. 2.6e-94;
Matches 295; Conservative 21; Mismatches 14; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 61

Qy 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 121

Qy 121 TGEQALEICDALARGAVDVIVDSVAALTTPKAEIEGEIGDSHMGGLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALTRSGAVDVIVDSVAALTTPKAEIEGEIGDSHMGGLAARMMSQAMRKLKAG 181

Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
Db 182 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 241

Qy 241 ETRVKVKNKTAAPKQAEFQILYCEGINFYGELVDLGKVKLEKAGAWYSYKGEKIQG 300
Db 242 ETRVKVKNKTAAPKQAEFQILYCEGINFYGELVDLGKVKLEKAGAWYSYKGEKIQG 301

Qy 301 GKANATAWLKNPETAKEIEKKVRELLLSN 330
Db 302 GKANATVLYKEHPMEYNELNKLREMLNH 331

RESULT 9
B82310
recA protein VC0543 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: B82310
R;Heideberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: B82310

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-412 <HEI>

A:Cross-references: UNIPARC:UPI0000164B55; GB:AE004140; GB:AE003852; NID:g9654965; PIDN:

A:Experimental source: serogroup O1; strain N16961; biotype E1 Tor

C:Genetics:

A:Gene: VC0543

A:Map position: 1

C:Superfamily: recombination protein recA

Query Match 85.6%; Score 1436; DB 2; Length 412;

Best Local Similarity 84.3%; Pred. No. 1.2e-89;

Matches 279; Conservative 31; Mismatches 21; Indels 0; Gaps 0;

Qy 2 IDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLDIALGAGGLPMGRI 61

Db 59 MDENKQKALAAALGQIEKQFGKGSIMRLGDNRMVDVETISTGSLDIALGAGGLPMGRI 118

Qy 62 VEIYGPSSGKTTTLTQVIAAAQREGKTCFADIAEHALDPIYARKLGVDIDNLLCSQDPT 121

Db 119 VEIYGPSSGKTTTLTLEIAAAQREGKTCFADIAEHALDPYAKLGWNIDELLVSQDPT 178

Qy 122 GEOALEICDALARGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGLAARMMSQAMRKLAGN 181

Db 179 GEOALEICDALARGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGLOARMLSQAMRKLTGN 238

Qy 182 LKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGSE 241

Db 239 LKQSNCMCIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRTGAIKEGEDEVVNE 298

Qy 242 TRVKVKNKIAAPPKQAEFOILYGEINFGYELVDLGVKEKLEIKAGAWYSYKGEKIGQG 301

Db 299 TRIKVKNKIAAPKAEANTQIMYGQGFNREGELDLGVKHKWVEKSGAWSYNGDKIGQG 358

Qy 302 KANATAMLKONPETAKEIEKKVRELLLSNP 332

Db 359 KANAYKYLKENPEIAKTLDKKREMLLNPN 389

RESULT 10

S46274

recombination protein recA - Vibrio cholerae

N:Alternate names: recombinase A

C:Species: Vibrio cholerae

C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 02-Feb-2001

C:Accession: S46274

R:Stroehner, U.H.; Lech, A.J.; Manning, P.A.

Mol. Gen. Genet. 244, 295-302, 1994

A:Title: Gene sequence of recA(+) and construction of recA mutants of Vibrio cholerae.

A:Reference number: S46274; MUID:94335880; PMID:8058040

A:Accession: S46274

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-354 <STR>

A:Cross-references: UNIPARC:UPI0000170586; EMBL:X71969; NID:g530269; PIDN:CAA50764.1; PI

C:Genetics:

A:Gene: recA

C:Superfamily: recombination protein recA

C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop;

F:65-72/Region: nucleotide-binding motif A (P-loop)

F:139-144/Region: nucleotide-binding motif B

F:71/Binding site: ATP (lys) #status predicted

Query Match 84.4%; Score 1417; DB 2; Length 354;

Best Local Similarity 83.1%; Pred. No. 1.9e-88;

Matches 275; Conservative 31; Mismatches 25; Indels 0; Gaps 0;

Qy 2 IDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLDIALGAGGLPMGRI 61

Db 1 MDENKQKALAAALGQIEKQFGKGSIMHLDNRMVDVETISTGSLDITLGGSLPMGRI 60

Qy 62 VEIYGPSSGKTTTLTQVIAAAQREGKTCFADIAEHALDPIYARKLGVDIDNLLCSQDPT 121

Db 61 VEIYGPSSGKTTTLTLEIAAAQREGKTCFADITEHALDPYAKLGWNIDELLVSQDPT 120

Qy 122 GEOALEICDALARGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGLAARMMSQAMRKLAGN 181

Db 121 GEOALEICDALARGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGLOARMLSQAMRKLTGN 180

Qy 182 LKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGSE 241

Db 181 LKQSNCMCIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRTGAIKEGEDEVVNE 240

Qy 242 TRVKVKNKIAAPPKQAEFOILYGEINFGYELVDLGVKEKLEIKAGAWYSYKGEKIGQG 301

Db 241 TRIKVVKNKIAAPKAEANTQIMYGQGFNREGELDLGVKHKWVEKSGAWSYNGDKIGQG 300

Qy 302 KANATAMLKONPETAKEIEKKVRELLLSNP 332

Db 301 KANACKYKLENPEIAKTLDKKREMLLNPN 331

RESULT 11

QJ1461

recombination protein recA - Vibrio anguillarum

N:Alternate names: recombinase A

C:Species: Vibrio anguillarum

C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004

C:Accession: QJ1461

R:Tolmashy, M.B.; Gammie, A.E.; Crosa, J.H.

Gene 110, 41-48, 1992

A:Title: Characterization of the recA gene of Vibrio anguillarum.

A:Reference number: QJ1461; MUID:92184113; PMID:1544576

A:Accession: QJ1461

A:Molecule type: DNA

A:Residues: 1-348 <TOL>

A:Cross-references: UNIPROT:P26348; UNIPARC:UPI0000133544; GB:M80525; NID:gl55256; PIDN

A:Experimental source: strains 775 and 531a

C:Genetics:

A:Gene: recA

C:Function:

C:Superfamily: recombination protein recA

C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop

F:65-72/Region: nucleotide-binding motif A (P-loop)

F:139-144/Region: nucleotide-binding motif B

F:256-279/Region: ATP binding #status predicted

F:71/Binding site: ATP (lys) #status predicted

Query Match 84.2%; Score 1412.5; DB 2; Length 348;

Best Local Similarity 83.6%; Pred. No. 3.8e-88;

Matches 276; Conservative 30; Mismatches 23; Indels 1; Gaps 1;

Qy 2 IDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLDIALGAGGLPMGRI 61

Db 1 MDENKQKALAAALGQIEKQFGKGSIMRLGDNRTMDVETISTGSLDIALGAGGLPMGRI 60

Qy 62 VEIYGPSSGKTTTLTQVIAAAQREGKTCFADIAEHALDPIYARKLGVDIDNLLCSQDPT 121

Db 61 VEIYGPSSGKTTTLTLEIAAAQREGKTCFADIAEHALDPYAKLGWNIDELLVSQDPT 120

Qy 122 GEOALEICDALARGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGLAARMMSQAMRKLAGN 181

Db 121 GEOALEICDALARGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGLOARMLSQAMRKLTGN 180

Qy 182 LKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGSE 241

Db 181 LKQSNCMCIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRTGAIKEGEDEVVNE 240

Qy 242 TRVKVKNKIAAPPKQAEFOILYGEINFGYELVDLGVKEKLEIKAGAWYSYKGEKIGQG 301

Db 241 TRIKVVKNKIAAPKQADTQILYGGQFNREGELVDLGVKHKLVEKAGAWSYNGDKIGQG 300

Qy 302 KANATAMLKONPETAKEIEKKVRELLLSNP 331

||||| :|:| | :|:| | | :|:| | | |

Db 301 KANACKFLRENPAAAMALDTKLEMLL-NP 329

RESULT 12

A49929

recombination protein recA - Haemophilus influenzae (strain Rd KW20)

N:Alternate names: rec-1 protein; recombinase A

C:Species: Haemophilus influenzae

C>Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 09-Jul-2004

C:Accession: A49929; A64080

R:Zulty, J.J.; Barcak, G.J.

J: Bacteriol. 175, 7269-7281, 1993

A:Title: Structural organization, nucleotide sequence, and regulation of the Haemophilus influenzae recombination protein

A:Reference number: A49929; MUID:94042901; PMID:8228674

A:Accession: A49929

A:Molecule type: DNA

A:Residues: 1-354 <ZUL>

A:Cross-references: UNIPROT:P43705; UNIPARC:UPI00001334E8; GB:L07529; NID:G305381; PIDN:A49929

A:Experimental source: Rd strain KW20

R:Flerschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. A:Reference number: A64000; MUID:95350630; PMID:7542800

A:Accession: A64080

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-354 <TIGR>

A:Cross-references: UNIPARC:UPI00001334E8; GB:U32741; GB:L42023; NID:G1573582; PIDN:AAC2

A:Experimental source: Rd strain KW20

C:Genetics:

A:Gene: recA

C:Function:

A:Description: plays an essential role in homologous recombination, in induction of the C:Superfamily: recombination protein recA

C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop; F:67-74/Region: nucleotide-binding motif A (P-loop)

F:141-146/Region: nucleotide-binding motif B

F:73/Binding site: ATP (lys) #status predicted

Query Match 79.4%; Score 1332; DB 2; Length 354;

Best Local Similarity 76.4%; Pred. No. 1.1e-82;

Matches 253; Conservative 42; Mismatches 36; Indels 0; Gaps 0;

Qy 1 ATDENKQKALAAALGOIEKQFGKSGIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 60

Db 2 ATQEQKQKALAAALGOIEKQFGKSGIMKLGDTKTLTLDVTSISTGSLGLDVALGIGGLPMGR 61

Qy 61 IVEIYGPSSSGKTTTLTQVIAAQAQKGTCAFDIAEHALDPIYARKLGVVDINLLCSQPD 120

Db 62 IVEIYGPSSSGKTTTLTSLVIAQAQKGTCAFDIAEHALDPIYARKLGVVDKELFVSQPD 121

Qy 121 TGEQALEICDALARGAVDVIIVDSVAALTPKAEIEGIDSHMGLAARMWSQAMRKLAG 180

Db 122 NGEQALEICDALVRSGAIDVIIVDSVAALTPKAEIEGDMGDSHMGLOARLMSQALRKLTG 181

Qy 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVYGS 240

Db 182 QIKNANCLVVFVFNQIRMKIGVMFGNPETTTGGNALKFYSSVRLDIRRTGSSVKDGENIIGN 241

Qy 241 ETRVKVKNKIAAPKQAEFQILYEGEINFYGELVDLGVKEKLI EKAGAWYSYKGEKIQG 300

Db 242 ETRVKVKNKLAAPPRQVDFQILYEGEISKAGELLELGVGKLVKESGAWTSYNGEKIQG 301

Qy 301 GKANATAMLKONPETAKEIEKKVRELLISNP 331

Db 302 GKANSKMKWLNENIEKSDLEARLRAELVANP 332

RESULT 13

JC5198

recombination protein recA - Aeromonas salmonicida

N:Alternate names: recombinase A

C:Species: Aeromonas salmonicida

C>Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004

C:Accession: JC5198

R:Umelo, E.; Noonan, B.; Trust, T.J.

Gene 175, 133-136, 1996

A:Title: Cloning, characterization and expression of the recA gene of Aeromonas salmonicida

A:Reference number: JC5198; MUID:97074662; PMID:8917089

A:Accession: JC5198

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-353 <UME>

A:Cross-references: UNIPROT:P94190; UNIPARC:UPI00001334C5; GB:U83688; NID:G1785949; PIDN:JC5198

C:Genetics:

A:Gene: recA

C:Function:

A:Description: plays an essential role in homologous recombination, in induction of the C:Superfamily: recombination protein recA

C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop; F:65-72/Region: nucleotide-binding motif A (P-loop)

F:139-144/Region: nucleotide-binding motif B

F:71/Binding site: ATP (lys) #status predicted

Query Match 78.8%; Score 1323; DB 2; Length 353;

Best Local Similarity 79.3%; Pred. No. 4.3e-82;

Matches 260; Conservative 30; Mismatches 38; Indels 0; Gaps 0;

Qy 2 IDENKQKALAAALGOIEKQFGKSGIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGRI 61

Db 1 MDQNKQKALAAALGOIEKQFGKSGIMLLGDSKTDWIDEAISTGSLSLDVALGIGGLPCGRI 60

Qy 62 VEIYGPSSSGKTTTLTQVIAAQAQREGKTCAPIDAEHALDPIYARKLGVVDINLLCSQPD 121

Db 61 VEIYGPSSSGKTTTLTQVIAEQAQKGVKCAPIDAEHALDPIYARKLGVNVVDLLISQSDT 120

Qy 122 GEQALEICDALARGAVDVIIVDSVAALTPKAEIEGIDSHMGLAARMWSQAMRKLAGN 181

Db 121 GEQALEICDMLVRNSAVDVIIVDSVAALTPKAEIEGEGDSHVGLQARLMSQALRKLTAN 180

Qy 182 LKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVYGS 241

Db 181 IKNANCLCIFINQIRMKIGVMFGSPETTTGGNALKFYASVRLDIRRIGAIKEGDEVVNGE 240

Qy 242 TRVKVKNKIAAPKQAEFQILYEGEINFYGELVDLGVKEKLI EKAGAWYSYKGEKIQG 301

Db 241 TRVKVKNKIVAPPKQAEFQIFYGVGISKEGELVDLGVKHKLIDKAGAWYSYNGEKIQG 300

Qy 302 KANATAMLKONPETAKEIEKKVRELLIS 329

Db 301 KANVMKLFTEKNVMAAEVEARLRELLIS 328

RESULT 14

RQPSAA

RecA protein PA3617 [imported] - Pseudomonas aeruginosa (strain PAO1)

N:Alternate names: recombinase A

C:Species: Pseudomonas aeruginosa

C>Date: 30-Sep-1990 #sequence_revision 30-Sep-1990 #text_change 09-Jul-2004

C:Accession: S06265; S10458; A49854; JN0304; G83192

R:Sano, Y.; Kageyama, M.

Mol. Gen. Genet. 208, 412-419, 1987

A:Title: The sequence and function of the recA gene and its protein in Pseudomonas aeruginosa

A:Reference number: S06265; MUID:88038334; PMID:2823059

A:Accession: S06265

A:Molecule type: DNA

A:Residues: 1-346 <SAN1>

A:Cross-references: UNIPROT:P08280; UNIPARC:UPI0000133516; GB:X05691; NID:G45413; PIDN:R:Zaitsev, E.N.; Krjukov, V.M.; Kuzmin, N.P.; Alekseev, A.A.; Lanzov, V.A. submitted to the EMBL Data Library, March 1990

A:Description: Restriction polymorphism and nucleotide sequence substitutions in the rec

A:Reference number: S10458
A:Accession: S10458
A:Molecule type: DNA
A:Residues: 1-346 <ZAI>
A:CROSS-references: UNIPARC:UPI0000133516; EMBL:X52261; NID:g45383; PIDN:CAA36504.1; PID:R.Sano, Y.
J. Bacteriol. 175, 2451-2454, 1993
A:Title: Role of the recA-related gene adjacent to the recA gene in *Pseudomonas aeruginosa*
A:Reference number: A49854; MUID:93224470; PMID:8468303
A:Accession: A49854
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 313-346 <SAN2>
A:CROSS-references: UNIPARC:UPI00001748B1; GB:D13090; NID:g286185
A:Note: sequence extracted from NCBI backbone (NCBI:129317, NCBI:P:129318)
R:Kryukov, V.M.; Zaitsev, E.N.; Kouzmin, N.P.; Bayev, A.A.
Bioorg. Khim. 16, 1177-1182, 1990
A:Title: Structure of the recA gene from *Pseudomonas aeruginosa*.
A:Reference number: JN0304; MUID:91182156; PMID:2127886
A:Accession: JN0304
A:Molecule type: DNA
A:Residues: 1-332, 'L', 334-346 <KRY>
A:CROSS-references: UNIPARC:UPI00001748B2; GB:X52261; NID:g45383; PIDN:CAA36504.1; PID:9
A:Experimental source: strain PAW 7
A:Note: the authors translated the codon CTG for residue 333 as Val
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: G83192
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-346 <STO>
A:CROSS-references: UNIPARC:UPI0000133516; GB:AE004782; GB:AE004091; NID:g9949772; PIDN:
A:Experimental source: strain PA01
C:Genetics:
A:Gene: recA; PA3617
C:Function:
A:Description: plays an essential role in homologous recombination, in induction of the
C:Superfamily: recombination protein recA
C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop;
F:65-72/Region: nucleotide-binding motif A (P-loop)
F:139-144/Region: nucleotide-binding motif B
F:71/Binding site: ATP (lys) #status predicted
Query Match 75.6%; Score 1268; DB 1; Length 346;
Best Local Similarity 74.4%; Pred. No. 2.1e-78;
Matches 244; Conservative 42; Mismatches 42; Indels 0; Gaps 0;
Qy 2 IDENKQKALAAALGOIEKQKFGKSGIMRLGEDRSMVDVETISGSLDIALGAGGLPMGRI 61
Db 1 MDENKKALAAALGOIEKQKFGKSGIMRLGEDRSMVDVETISGSLDIALGAGGLPMGRI 60
Qy 62 VEIYGPSSGKTTTLTLOVIAAAQREGKTCAPIDAEHALDPTIYARKLGVDIDNLLCSQPD 121
Db 61 VEIYGPSSGKTTTLTLOVIAAAQREGKTCAPIDAEHALDPTIYARKLGVDIDNLLCSQPD 120
Qy 122 GEOALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGIAARMMSQAMRKLAGN 181
Db 121 GEOALEITDMLVRNAVDVIIVDSVAALTPKAEIEGEIGDSHMGIAARMMSQAMRKLAGN 180
Qy 182 LKQSNLTLLIFNQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRGAVKEGENVVGSE 241
Db 181 IKNANCLVIFNQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRGAVKEGENVVGSE 240
Qy 242 TRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVEKLEIKAGAWSYKGEKIGQG 301
Db 241 TRVKVKNKVPFPQAEFQILYKGIYRNGEIIDLGVLGLVEKSGAWSYQSKIGQG 300
Qy 302 KANATAMLKONPETAKEIEKKVRELLS 329
Db 302 KANATAMLKONPETAKEIEKKVRELLS 329

Db 301 KANAAKYLEDNPEIGSVLEKTIIRDOLLA 328
RESULT 15
JN0321
recombination protein recA - *Azotobacter vinelandii*
N:Alternate names: recombinase A
C:Species: *Azotobacter vinelandii*
C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 09-Jul-2004
C:Accession: JN0321
R:Kenkatesh, T.V.; Das, H.K.
Gene 113, 47-53, 1992
A:Title: The *Azotobacter vinelandii* recA gene: sequence analysis and regulation of exp.
A:Reference number: JN0321; MUID:92225347; PMID:1563632
A:Accession: JN0321
A:Molecule type: DNA
A:Residues: 1-349 <KEN>
A:CROSS-references: UNIPROT:P29246; UNIPARC:UPI00001334CC
A:Note: the authors translated the codon GAC for residue 129 as Val and AAC for residu
C:Genetics:
A:Gene: recA
C:Function:
A:Description: plays an essential role in homologous recombination, in induction of thr
C:Superfamily: recombination protein recA
C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; DNA replication; nucleoti
F:65-72/Region: nucleotide-binding motif A (P-loop)
F:139-144/Region: nucleotide-binding motif B
F:71/Binding site: ATP (lys) #status predicted
Query Match 75.4%; Score 1265; DB 2; Length 349;
Best Local Similarity 73.5%; Pred. No. 3.5e-78;
Matches 244; Conservative 41; Mismatches 47; Indels 0; Gaps 0;
Qy 2 IDENKQKALAAALGOIEKQKFGKSGIMRLGEDRSMVDVETISGSLDIALGAGGLPMGRI 61
Db 1 MDENKKALAAALGOIEKQKFGKSGIMRLGEDRSMVDVETISGSLDIALGAGGLPMGRI 60
Qy 62 VEIYGPSSGKTTTLTLOVIAAAQREGKTCAPIDAEHALDPTIYARKLGVDIDNLLCSQPD 121
Db 61 VEIYGPSSGKTTTLTLOVIAAAQREGKTCAPIDAEHALDPTIYARKLGVDIDNLLCSQPD 120
Qy 122 GEOALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGIAARMMSQAMRKLAGN 181
Db 121 GEOALEITDMLVRNAVDVIIVDSVAALTPKAEIEGEIGDSHMGIAARMMSQAMRKLAGN 180
Qy 182 LKQSNLTLLIFNQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRGAVKEGENVVGSE 241
Db 181 IQNANCLVIFNQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRGAVKEGENVVGSE 240
Qy 242 TRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVEKLEIKAGAWSYKGEKIGQG 301
Db 241 TRVKVKNKVPFPQAEFQILYKGIYRNGEIIDLGVLGLVEKSGAWSYQSKIGQG 300
Qy 302 KANATAMLKONPETAKEIEKKVRELLSNPNS 333
Db 301 KANAAKPLEDNPEVAAAVEKSIKRDQLLAAPAS 332
Search completed: February 16, 2006, 01:24:56
Job time : 26.5 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 01:14:35 ; Search time 152 Seconds
(without alignments)
1554.947 Million cell updates/sec

Title: US-10-733-782-1

Perfect score: 1678

Sequence: 1 AIDENKQKALAAALGQIEKQ.....AKEIEKKVRELLSNPNSTP 335

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1678	100.0	352	1 RECA_ECO57	P0A798 escherichia
2	1678	100.0	352	1 RECA_ECOL6	P0A797 escherichia
3	1678	100.0	352	1 RECA_ECOL1	P0A796 escherichia
4	1678	100.0	352	1 RECA_SHIFL	P0A799 shigella fl
5	1671	99.6	352	1 RECA_SHISO	Q95ff6 shigella so
6	1641	97.8	352	1 RECA_SALTY	P65978 salmonella
7	1641	97.8	352	1 RECA_SALTY	P65977 salmonella
8	1641	97.8	352	2 Q57KU4_SALCH	Q57ku4 salmonella
9	1639	97.7	353	2 Q5PF15_SALPA	Q5pf15 salmonella
10	1547	92.2	355	1 RECA_PHOLL	Q7n7a6 photorhabdu
11	1537	91.6	357	1 RECA_ERWCT	Q6dl88 erwinia car
12	1534	91.4	355	1 RECA_YERPE	P37858 yersinia pe
13	1534	91.4	356	1 RECA_YERPS	P66e70 yersinia ps
14	1532	91.3	353	1 RECA_ENTAG	P33037 enterobacte
15	1527	91.0	353	1 RECA_SERMA	P17479 serratia ma
16	1524	90.8	358	1 RECA_XENBV	P96185 xenorhabdu
17	1523	90.8	355	1 RECA_SODGL	P62220 sodalis glo
18	1513	90.2	355	2 Q5Y131_9ENTR	Q5y131 primary end
19	1504	89.6	354	1 RECA_PROMI	P11406 proteus mir
20	1501	89.5	342	1 RECA_ERWCA	P26344 erwinia car
21	1501	89.5	357	1 RECA_SHEON	Q8eb50 shewanella
22	1500	89.4	358	1 RECA_XENNE	Q9X5p5 xenorhabdu
23	1486	88.6	325	1 RECA_PROVU	P26346 proteus vul
24	1437	85.6	354	1 RECA_VIBCH	P45383 vibrio chol
25	1429	85.2	347	1 RECA_VIBPA	Q871r1 vibrio para
26	1418	84.5	348	1 RECA_VIBNA	Q6xz07 vibrio natr
27	1418	84.5	349	1 RECA_VIBVU	Q8dc51 vibrio vuln
28	1418	84.5	349	1 RECA_VIBVU	Q7mhr4 vibrio vuln
29	1416	84.4	348	2 Q5E7G6_VIBFI	Q5e7g6 vibrio fisc
30	1412.5	84.2	348	1 RECA_VIBAN	P26348 vibrio angu
31	1404	83.7	352	1 RECA_PHOPR	Q61mu2 photobacter

32	1400	83.4	346	2 Q9S4R6_VIBCH	Q9s4r6 vibrio chol
33	1400	83.4	346	2 Q9R2W9_VIBCH	Q9r2w9 vibrio chol
34	1353	80.6	354	1 RECA_PASMU	P95526 pasteurella
35	1332	79.4	354	1 RECA_HAEIN	P43705 haemophilus
36	1332	79.4	354	2 Q4QMV2_HAEI8	Q4qmv2 haemophilus
37	1326	79.0	348	1 RECA_PSEOL	Q9r9u2 pseudomonas
38	1326	79.0	372	2 Q5QUB8_IDILO	Q5qub8 idiomarina
39	1323	78.8	353	1 RECA_AERSA	P94190 aeromonas s
40	1319	78.6	351	1 RECA_MANSM	Q65qb0 mannheimia
41	1317	78.5	331	1 RECA_WIGRR	Q8d2w7 wigglewort
42	1316	78.4	352	1 RECA_ACTAC	Q9jrp9 actinobacill
43	1288	76.8	368	1 RECA_PASHA	Q9rny0 pasteurella
44	1276	76.0	347	2 Q6EV36_PSEST	Q6ev36 pseudomonas
45	1275	76.0	349	2 Q4IXX9_AZOVI	Q4ixx9 azotobacter

ALIGNMENTS

RESULT 1

RECA_ECO57 STANDARD; PRT; 352 AA.
AC P0A7G8; P03017; P26347; P78213;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RECA protein (Recombinase A).
GN Name=reca; OrderedLocusNames=z4002, EC83556;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDU933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
NA Nature 409:529-533(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
NA DNA Res. 8:11-22(2001).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
single-stranded DNA, the ATP-dependent uptake of single-stranded
DNA by duplex DNA, and the ATP-dependent hybridization of
homologous single-stranded DNAs. It interacts with lexA causing
its activation and leading to its autocatalytic cleavage (By
similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the reca family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC EMBL; AE005174; AG57804.1; -; Genomic DNA.
CC DR EMBL; BA000007; BAB36979.1; -; Genomic DNA.
CC PIR; D91073; D91073.

```

DR HAMAP; MF_00268; -; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PR00142; RecA.
DR ProDom; PD000229; RecA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS50162; RECA_2; 1.
DR PROSITE; PS50163; RECA_3; 1.
KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
FT INIT_MET 0 By similarity.
FT NP_BIND 66 73 ATP (By similarity).
FT CONFLICT 139 139 D -> N (in Ref. 1).
SQ SEQUENCE 352 AA; 37842 MW; 989B02378EC4A402 CRC64;

Query Match 100.0%; Score 1678; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 5.4e-104;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 60
Db 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 60
Qy 61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAPDAEHALDPIYARKLGVVDINLLCSQPD 120
Dl 61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAPDAEHALDPIYARKLGVVDINLLCSQPD 120
Qy 121 TGEQALEICDALARGAVDVIWDSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 180
Db 121 TGEQALEICDALARGAVDVIWDSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 180
Qy 181 NLKQSNLTLLIFINQIRMKIGVFGNPGTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 181 NLKQSNLTLLIFINQIRMKIGVFGNPGTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Qy 241 ETRVKVKNKIAAPFKQAEFQILYEGEINFGYELVDLGVKEKLIIEKAGAWYSYKGEKIQG 300
Db 241 ETRVKVKNKIAAPFKQAEFQILYEGEINFGYELVDLGVKEKLIIEKAGAWYSYKGEKIQG 300
Qy 301 GKANATAWLKONPETAKIEKKVRELLSNPNSTP 335
Db 301 GKANATAWLKONPETAKIEKKVRELLSNPNSTP 335

```

RESULT 2

```

RECA_ECOLI6 STANDARD; PRT; 352 AA.
AC POA7G6; P03017; P26347; P78213;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocusNames=c3253;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.F., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded

```

```

CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with LexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the RecA family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AE016765; AA081704.1; -; Genomic_DNA.
CC HAMAP; MF_00268; -; 1.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR001553; RecA.
CC Pfam; PF00154; RecA; 1.
CC PRINTS; PR00142; RecA.
CC ProDom; PD000229; RecA; 1.
CC SMART; SM00382; AAA; 1.
CC TIGRFAMs; TIGR02012; tigrfam_recA; 1.
CC PROSITE; PS00321; RECA_1; 1.
CC PROSITE; PS50162; RECA_2; 1.
CC PROSITE; PS50163; RECA_3; 1.
KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
FT INIT_MET 0 By similarity.
FT NP_BIND 66 73 ATP (By similarity).
FT CONFLICT 139 139 D -> N (in Ref. 1).
SQ SEQUENCE 352 AA; 37842 MW; 989B02378EC4A402 CRC64;

Query Match 100.0%; Score 1678; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 5.4e-104;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 60
Db 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 60
Qy 61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAPDAEHALDPIYARKLGVVDINLLCSQPD 120
Db 61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAPDAEHALDPIYARKLGVVDINLLCSQPD 120
Qy 121 TGEQALEICDALARGAVDVIWDSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 180
Db 121 TGEQALEICDALARGAVDVIWDSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 180
Qy 181 NLKQSNLTLLIFINQIRMKIGVFGNPGTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 181 NLKQSNLTLLIFINQIRMKIGVFGNPGTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Qy 241 ETRVKVKNKIAAPFKQAEFQILYEGEINFGYELVDLGVKEKLIIEKAGAWYSYKGEKIQG 300
Db 241 ETRVKVKNKIAAPFKQAEFQILYEGEINFGYELVDLGVKEKLIIEKAGAWYSYKGEKIQG 300
Qy 301 GKANATAWLKONPETAKIEKKVRELLSNPNSTP 335
Db 301 GKANATAWLKONPETAKIEKKVRELLSNPNSTP 335

RESULT 3
RECA_ECOLI STANDARD; PRT; 352 AA.
AC POA7G6; P03017; P26347; P78213;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; Synonym=lexB; recH, rnmB, tif, umuB, zab;
OS OrderedLocusNames=b2699;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.

```

OX NCBI_TaxID=562;
 RN [1]
 RP NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 1-5.
 RX MEDLINE=80145618; PubMed=6244554;
 RA Horii T., Ogawa T., Ogawa H.;
 RT "Organization of the recA gene of *Escherichia coli*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 77:313-317(1980).
 RN [2]
 RP NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 1-2.
 RX MEDLINE=80234673; PubMed=6930655;
 RA Sancar A., Stachelle C., Konigsberg W., Rupp W.D.;
 RT "Sequences of the recA gene and protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 77:2611-2615(1980).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=91109725; PubMed=2274037;
 RA Zhao X.J., McEntee K.;
 RT "DNA sequence analysis of the recA genes from *Proteus vulgaris*,
 RT *Erwinia carotovora*, *Shigella flexneri* and *Escherichia coli* B/r.";
 RL Mol. Gen. Genet. 222:369-376(1990).
 RN [4]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Valdes J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=K12;
 RX MEDLINE=97349980; PubMed=9205837;
 RA Yamamoto Y., Aiba H., Baba T., Hayaishi K., Inada T., Isono K.,
 RA Itch T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
 RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,
 RA Oshima T., Oyama S., Saito N., Saito N., Saito Y., Sivaundaram S.,
 RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
 RA Yamagata S., Horiuchi T.;
 RT "Construction of a contiguous 874-kb sequence of the *Escherichia coli*-
 RT K12 genome corresponding to 50.0-68.8 min on the linkage map and
 RT analysis of its sequence features.";
 RL DNA Res. 4:91-113(1997).
 RN [6]
 RP PROTEIN SEQUENCE OF 89-106 AND 178-183.
 RX MEDLINE=9525284; PubMed=7737176;
 RA Morimatsu K., Horii T.;
 RT "The DNA-binding site of the RecA protein. Photochemical cross-linking
 RT of Tyr103 to single-stranded DNA.";
 RL Eur. J. Biochem. 228:772-778(1995).
 RN [7]
 RP PROTEIN SEQUENCE OF 186-193.
 RX MEDLINE=96067680; PubMed=7588783;
 RA Gardner R.V., Voloshin O.N., Camerini-Otero R.D.;
 RT "The identification of the single-stranded DNA-binding domain of the
 RT *Escherichia coli* RecA protein.";
 RL Eur. J. Biochem. 233:419-425(1995).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE=92114994; PubMed=1731246; DOI=10.1038/355318a0;
 RA Story R.M., Weber I.T., Steitz T.A.;
 RT "The structure of the *E. coli* recA protein monomer and polymer.";
 RL Nature 355:318-325(1992).
 RN [9]
 RP ERRATUM.
 RA Story R.M., Weber I.T., Steitz T.A.;
 RL Nature 355:567-567(1992).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE=97185905; PubMed=9033586;
 RA Xu X., Egelman E.H.;
 RT "The RecA hexamer is a structural homologue of ring helicases.";

RL Nat. Struct. Biol. 4:101-104(1997).
 RN [11]
 RP STRUCTURE OF ATP-BINDING FOLD.
 RX MEDLINE=92115005; PubMed=1731253; DOI=10.1038/355374a0;
 RA Story R.M., Steitz T.A.;
 RT "Structure of the recA protein-ADP complex.";
 RL Nature 355:374-376(1992).
 CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
 CC single-stranded DNA, the ATP-dependent uptake of single-stranded
 CC DNA by duplex DNA, and the ATP-dependent hybridization of
 CC homologous single-stranded DNAs. It interacts with LexA causing
 CC its activation and leading to its autocatalytic cleavage.
 CC -!- INTERACTION:
 CC P0A6G7:clpP; NbExp=1; IntAct=EBI-370331, EBI-370625;
 CC P08622:dnaJ; NbExp=1; IntAct=EBI-370331, EBI-545285;
 CC P77488:dxs; NbExp=1; IntAct=EBI-370331, EBI-549927;
 CC P06138:ftsZ; NbExp=1; IntAct=EBI-370331, EBI-370963;
 CC P31120:glmM; NbExp=1; IntAct=EBI-370331, EBI-370683;
 CC P09097:gyrA; NbExp=1; IntAct=EBI-370331, EBI-547129;
 CC P0A6H5:halU; NbExp=1; IntAct=EBI-370331, EBI-369317;
 CC P0A817:metK; NbExp=1; IntAct=EBI-370331, EBI-546295;
 CC P13519:metK; NbExp=1; IntAct=EBI-370331, EBI-371008;
 CC P09152:narG; NbExp=1; IntAct=EBI-370331, EBI-547248;
 CC P05055:ppp; NbExp=1; IntAct=EBI-370331, EBI-548080;
 CC P0A7B9:pyrH; NbExp=1; IntAct=EBI-370331, EBI-370182;
 CC P03842:rph; NbExp=1; IntAct=EBI-370331, EBI-557453;
 CC P36979:yfgB; NbExp=1; IntAct=EBI-370331, EBI-559071;
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- INDUCTION: In response to low temperature. Sensitive to
 CC temperature through changes in the linking number of the DNA.
 CC -!- SIMILARITY: Belongs to the recA family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; V00328; CAA23618.1; -; Genomic DNA.
 CC EMBL; U00096; AAC75741.1; -; Genomic DNA.
 CC EMBL; D90892; BAA15561.1; -; Genomic DNA.
 CC PIR; G65049; RQBCA.
 CC PDB; 1AA3; NMR; @=-.
 CC PDB; 1N03; EM; A/B/C/D/E/F/G=1-352.
 CC PDB; 1REA; X-ray; @=-.
 CC PDB; 1U94; X-ray; A=1-352.
 CC PDB; 1U98; X-ray; A=1-352.
 CC PDB; 1U99; X-ray; A=1-352.
 CC PDB; 1XMS; X-ray; A=1-352.
 CC PDB; 1XWV; X-ray; A=1-352.
 CC PDB; 2REB; X-ray; @=-.
 CC PDB; 2REC; EM; A/B/C/D/E/F=1-352.
 CC IntAct; P0A7G6; -; 1.
 CC SWISS-2DPAGE; P0A7G6; COLI.
 CC ECO2DBASE; C039.3; 6TH EDITION.
 CC EcoGene; E08016; -; 1.
 CC HAMAP; MF_00268; -; 1.
 CC InterPro; IPR003593; AAA_ATPase.
 CC InterPro; IPR001553; RecA.
 CC Pfam; PF00154; RecA; 1.
 CC PRINTS; PR00142; RECA.
 CC ProDom; PD000229; RecA; 1.
 CC SMART; SM00382; AAA; 1.
 CC TIGRFAMs; TIGR02012; tigrfam_recA; 1.
 CC PROSITE; PS00321; RECA_1; 1.
 CC PROSITE; PS0162; RECA_2; 1.
 CC PROSITE; PS0163; RECA_3; 1.
 CC 3D-structure; ATP-binding; Complete proteome;
 CC Direct protein sequencing; DNA damage; DNA recombination; DNA repair;
 CC DNA-binding; Nucleotide-binding; SOS response.
 FT INIT MET 0
 FT NP_BIND 66 73 ATP.

```
FT CONFLICT 112 112 D -> E (in Ref. 5).
FT TURN 190 190 Missing (in Ref. 7).
FT TURN 4 4
FT HELIX 5 21
FT HELIX 23 25
FT TURN 29 30
FT STRAND 39 40
FT HELIX 45 50
FT TURN 51 51
FT STRAND 56 57
FT TURN 58 59
FT STRAND 61 65
FT TURN 62 69
FT HELIX 72 85
FT TURN 86 87
FT STRAND 90 94
FT HELIX 101 106
FT TURN 107 108
FT HELIX 111 113
FT STRAND 115 117
FT HELIX 122 135
FT STRAND 140 144
FT HELIX 146 148
FT HELIX 152 155
FT TURN 166 185
FT TURN 186 186
FT STRAND 188 193
FT HELIX 213 218
FT STRAND 221 233
FT TURN 234 235
FT STRAND 236 249
FT STRAND 257 263
FT TURN 264 266
FT STRAND 267 268
FT HELIX 270 280
FT TURN 281 282
FT STRAND 285 287
FT TURN 288 289
FT STRAND 290 293
FT TURN 294 295
FT STRAND 296 300
FT HELIX 301 311
FT HELIX 313 327

Query Match 100.0%; Score 1678; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 5.4e-104;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPWGR 60
DB 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPWGR 60

QY 61 IVEIYGPSSGKTTLTLOVIAAQRGKTCFIDEHALDPIYARKLGVDIDNLLCSQPD 120
DB 61 IVEIYGPSSGKTTLTLOVIAAQRGKTCFIDEHALDPIYARKLGVDIDNLLCSQPD 120

QY 121 TGEQALEICDALARSGVDVIVVDSVAALTTPKAEIEGEIGDSHMGLAARMWSQAMKLAG 180
DB 121 TGEQALEICDALARSGVDVIVVDSVAALTTPKAEIEGEIGDSHMGLAARMWSQAMKLAG 180

QY 181 NLKQNTLLIFINQIRMKIGVFGNPETTTGGNALKFVASVRLDIRRGAVKEGNNVGS 240
DB 181 NLKQNTLLIFINQIRMKIGVFGNPETTTGGNALKFVASVRLDIRRGAVKEGNNVGS 240

QY 241 ETRVVKVKNKIAAPFKQAEFQLYGEINFGYELVDLGKVKELIEKAGAWSYKGEKIGQ 300
DB 241 ETRVVKVKNKIAAPFKQAEFQLYGEINFGYELVDLGKVKELIEKAGAWSYKGEKIGQ 300

QY 301 GKANATAMKONPETAKIEKKVRELLLSNPSTP 335
DB 301 GKANATAMKONPETAKIEKKVRELLLSNPSTP 335
```

```
RESULT 4
RECA_SHIFL STANDARD; PRT; 352 AA.
ID AC POA7G9; P03017; P26347; P78213;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocusNames=Sf2722, S2913;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BS12 / Serotype 2a;
RX MEDLINE=91109725; PubMed=2274037;
RA Zhao X. J., McEntee K. J.
RT "DNA sequence analysis of the recA genes from Proteus vulgaris,
RT Erwinia carotovora, Shigella flexneri and Escherichia coli B/r.";
RL Mol. Gen. Genet. 222:369-376(1990).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J. J.
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
DOI=10.1128/IAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M. B., Burland V., Venkatesan M. M., Deng W.,
RA Fournier G., Mayhew G. F., Plunkett G. III, Rose D. J., Darling A.,
RA Mau B., Perna N. T., Payne S. M., Runyen-Janecky L. J., Zhou S.,
RA Schwartz D. C., Blattner F. R. J.
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with LexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; X55553; CAB56806.1; -; Genomic DNA.
CC EMBL; AE005674; AAN44214.1; ALT_INIT; Genomic DNA.
CC EMBL; AE016987; AAP18040.1; -; Genomic DNA.
CC HAMAP; MF_00268; -; 1.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR001553; RecA.
CC Pfam; PF00154; RecA; 1.
CC PRINTS; PR00142; RecA.
CC ProDom; PD000229; RecA; 1.
CC SMART; SM00382; AAA; 1.
CC TIGRFAMs; TIGR02012; tigrfam_recA; 1.
CC PROSITE; PS00321; RECA_1; 1.
CC PROSITE; PS50162; RECA_2; 1.
```

```

DR PROSITE, PS50163; RECA_3; 1.
KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
DR INIT MET 0 By similarity.
FT NP BIND 66 73 ATP (By similarity).
SQ SEQUENCE 352 AA; 37842 MW; 989B02378EC4A402 CRC64;

Query Match
Best Local Similarity 100.0%; Score 1678; DB 1; Length 352;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSDMDVETISTGSLSLDIALGAGGLPMGR 60
DB 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSDMDVETISTGSLSLDIALGAGGLPMGR 60
QY 61 IVEIYGPESSGKTTLTQVIAAAREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 61 IVEIYGPESSGKTTLTQVIAAAREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
QY 121 TGEQALEICDALARSGADVIVDVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLK 180
DB 121 TGEQALEICDALARSGADVIVDVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLK 180
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
DB 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
QY 241 ETRVKVVKNTAAAPKQAEFQILYGEINFGYVGLVDLGVEKLEKAGAWYSYKGEKIQ 300
DB 241 ETRVKVVKNTAAAPKQAEFQILYGEINFGYVGLVDLGVEKLEKAGAWYSYKGEKIQ 300
QY 301 GKANATAWLKONPETAKEIEKKVRELLSNPNSTP 335
DB 301 GKANATAWLKONPETAKEIEKKVRELLSNPNSTP 335

RESULT 5
RECA_SHISO STANDARD; PRT; 352 AA.
AC Q92FF6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA;
OS Shigella sonnei.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=624;
[1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=KNH104S;
RA Park Y.C., Shin H.J., Kim Y.C.;
RT "Cloning and nucleotide sequence of the recA gene from Shigella sonnei
  KNH104S."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with LexA causing
CC its activation and leading to its autocatalytic cleavage.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AF101227; AAC72856.1; -; Genomic_DNA.
CC HSSP; P03017; 2REB.
CC SRR; Q92FF6; 3-328.

```

```

DR HAMAP; MF_00268; -; 1.
DR InterPro; IPR003593; AAA ATPase.
DR RepairPro; IPR001553; RecA.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PR00142; RECA.
DR PRODOM; PD000229; RecA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS0162; RECA_2; 1.
DR PROSITE; PS0163; RECA_3; 1.
KW ATP-binding; DNA damage; DNA recombination; DNA repair; DNA-binding;
KW Nucleotide-binding; SOS response.
DR INIT MET 0 By similarity.
FT NP BIND 66 73 ATP (By similarity).
SQ SEQUENCE 352 AA; 37799 MW; E5C50231893AA40C CRC64;

Query Match
Best Local Similarity 99.6%; Score 1671; DB 1; Length 352;
Matches 334; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSDMDVETISTGSLSLDIALGAGGLPMGR 60
DB 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSDMDVETISTGSLSLDIALGAGGLPMGR 60
QY 61 IVEIYGPESSGKTTLTQVIAAAREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 61 IVEIYGPESSGKTTLTQVIAAAREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
QY 121 TGEQALEICDALARSGADVIVDVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLK 180
DB 121 TGEQALEICDALARSGADVIVDVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLK 180
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
DB 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
QY 241 ETRVKVVKNTAAAPKQAEFQILYGEINFGYVGLVDLGVEKLEKAGAWYSYKGEKIQ 300
DB 241 ETRVKVVKNTAAAPKQAEFQILYGEINFGYVGLVDLGVEKLEKAGAWYSYKGEKIQ 300
QY 301 GKANATAWLKONPETAKEIEKKVRELLSNPNSTP 335
DB 301 GKANATAWLKONPETAKEIEKKVRELLSNPNSTP 335

RESULT 6
RECA_SALTI STANDARD; PRT; 352 AA.
AC P65978; O8XET0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocusNames=STV2950, t2730;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
[1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC MEDLINE=21534947; PubMed=1677608; DOI=10.1038/35101607;
RX STRAIN=CT18;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
  enterica serovar Typhi CT18."

```

```
RL Nature 413:848-852(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA DOI=10.1128/JB.185.7.2330-2337.2003;
RA Deng W., liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and C718.";
RL J. Bacteriol. 185:2330-2337(2003).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AL627276; CAD05935.1; -; Genomic_DNA.
CC EMBL; AE016843; AAO70291.1; -; Genomic_DNA.
CC HSSP; P03017; 2REB.
CC SMR; P65978; 3-328.
CC HAMAP; MF_00268; 1.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR001553; RecA.
CC Pfam; PF00154; RecA; 1.
CC PRINTS; PR00142; RecA.
CC ProDom; PD000229; RecA; 1.
CC SMART; SM00382; AAA; 1.
CC TIGRFAMs; TIGR02012; tigrfam_recA; 1.
CC PROSITE; PS00321; RECA_1; 1.
CC PROSITE; PS0162; RECA_2; 1.
CC PROSITE; PS0163; RECA_3; 1.
CC ATP-binding; Complete proteome; DNA damage; DNA recombination;
CC DNA repair; DNA-binding; Nucleotide-binding; SOS response.
CC INIT MET 0
CC NP_BIND 66 73 ATP (By similarity).
CC SEQUENCE 352 AA; 37813 MW; F24E51CB9A31B818 CRC64;

Query Match 97.8%; Score 1641; DB 1; Length 352;
Best Local Similarity 97.3%; Pred. No. 1.6e-101;
Matches 326; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
DB 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
QY 61 IVEIYGPSSGKTTTLTQVIAAAAREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 61 IVEIYGPSSGKTTTLTQVIAAAAREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120
QY 121 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGETGDSHMGLAARMMSQAMRKLAG 180
DB 121 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGETGDSHMGLAARMMSQAMRKLAG 180
QY 181 NLKQNTLLIFNQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVYGS 240
DB 181 NLKQNTLLIFNQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVYGS 240
QY 241 ETRVKVKNKIAAPFKQAFQILYEGEINFYGELVDLGKVEKLEIKAGAWYSGEKGIGQ 300
DB 241 ETRVKVKNKIAAPFKQAFQILYEGEINFYGELVDLGKVEKLEIKAGAWYSGEKGIGQ 300
QY 301 GKANATAWLKNDPETAKEIEKRVRELLLSNQNPSTP 335
DB 301 GKANATAWLKNDPETAKEIEKRVRELLLSNQNPSTP 335
```

```
Db 301 GKANATTWLKENPATAKEIEKRVRELLLSNQNPATP 335

RESULT 7
ID RECA_SALTY STANDARD; PRT; 352 AA.
AC P65977; OSXET0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocusNames=STM2829;
OS Salmonella typhimurium
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AE008829; AAL21709.1; -; Genomic_DNA.
CC HSSP; P03017; 2REB.
CC SMR; P65977; 3-328.
CC StyGene; SG7777; recA.
CC HAMAP; MF_00268; -; 1.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR001553; RecA.
CC Pfam; PF00154; RecA; 1.
CC PRINTS; PR00142; RecA.
CC ProDom; PD000229; RecA; 1.
CC SMART; SM00382; AAA; 1.
CC TIGRFAMs; TIGR02012; tigrfam_recA; 1.
CC PROSITE; PS00321; RECA_1; 1.
CC PROSITE; PS0162; RECA_2; 1.
CC PROSITE; PS0163; RECA_3; 1.
CC ATP-binding; Complete proteome; DNA damage; DNA recombination;
CC DNA repair; DNA-binding; Nucleotide-binding; SOS response.
CC INIT MET 0
CC NP_BIND 66 73 ATP (By similarity).
CC SEQUENCE 352 AA; 37813 MW; F24E51CB9A31B818 CRC64;

Query Match 97.8%; Score 1641; DB 1; Length 352;
Best Local Similarity 97.3%; Pred. No. 1.6e-101;
Matches 326; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
DB 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
QY 61 IVEIYGPSSGKTTTLTQVIAAAAREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120
```

Db 61 IVEIYGPSSGKTTTLQVIAAAREGKTCAFIDAEHALDPYARKLGVVDINLLCSQPD 120
QY 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGCLAARMMSQAMRKLKAG 180
Db 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGCLAARMMSQAMRKLKAG 180
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGDNVVGVS 240
Db 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGDNVVGVS 240
QY 241 ETRVKVVKNTAAAPFKQAEFOILYCEGINFYGELVDLGVKELIEKAGAWTSYKGEKIQG 300
Db 241 ETRVKVVKNTAAAPFKQAEFOILYCEGINFYGELVDLGVKELIEKAGAWTSYKGEKIQG 300
QY 301 GKANATATLWLNKPNPATAKEIEKKVRELLLSNPNSTP 335
Db 301 GKANATATLWLNKPNPATAKEIEKKVRELLLSNPNATP 335

RESULT 8
Q57KU4 SALCH
ID Q57KU4 SALCH PRELIMINARY; PRT; 353 AA.
AC Q57KU4
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE DNA strand exchange and recombination protein with protease and
nuclease activity.
GN Name=recA; OrderedLocusNames=SC2762;
OS Salmonella choleraesuis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=591;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SC-B67;
RX PubMed=15781495; DOI=10.1093/nar/gki297;
RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
Wang H.-S., Lee Y.-S.
RT "The genome sequence of Salmonella enterica serovar Choleraesuis, a
highly invasive and resistant zoonotic pathogen."
RL Nucleic Acids Res. 33:1690-1698(2005).
EMBL; AE017220; AAX66668.1; -; Genomic_DNA.
DR SMR; Q57KU4; 4-329.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003684; F:damaged DNA binding; IEA.
DR GO; GO:0008094; F:DNA-dependent ATPase activity; IEA.
DR GO; GO:0017111; F:nucleoside-triphosphatase activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA.
DR PRINTS; PR00142; RECA.
DR SMART; SMO0382; AAA.
DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS0162; RECA_2; 1.
DR PROSITE; PS0163; RECA_3; 1.
KW Complete proteome.
SQ SEQUENCE 353 AA; 37944 MW; ECBAB042F16362C CRC64;

Query Match 97.8%; Score 1641; DB 2; Length 353;
Best Local Similarity 97.3%; Pred. No. 1.6e-101;
Matches 326; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSDIALGAGGLPMGR 61
QY 61 IVEIYGPSSGKTTTLQVIAAAREGKTCAFIDAEHALDPYARKLGVVDINLLCSQPD 120

Db 62 IVEIYGPSSGKTTTLQVIAAAREGKTCAFIDAEHALDPYARKLGVVDINLLCSQPD 121
QY 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGCLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGCLAARMMSQAMRKLKAG 181
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGDNVVGVS 240
Db 182 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGDNVVGVS 241
QY 241 ETRVKVVKNTAAAPFKQAEFOILYCEGINFYGELVDLGVKELIEKAGAWTSYKGEKIQG 300
Db 242 ETRVKVVKNTAAAPFKQAEFOILYCEGINFYGELVDLGVKELIEKAGAWTSYKGEKIQG 301
QY 301 GKANATATLWLNKPNPATAKEIEKKVRELLLSNPNSTP 335
Db 302 GKANATATLWLNKPNPATAKEIEKKVRELLLSNPNATP 336

RESULT 9
Q5PF15 SALPA
ID Q5PF15 SALPA PRELIMINARY; PRT; 353 AA.
AC Q5PF15;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE RecA protein.
GN Name=recA; OrderedLocusNames=SPA2687;
OS Salmonella paratyphi-a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=54388;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 9150 / SARB42;
RX PubMed=15531882; DOI=10.1038/ngl470;
RA McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,
Porwollik S., Sabo A., Meyer R., Bieri T., Ozeraky P., McLellan M.,
Harkins C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
Kohlberg S., Strong C., Du F., Carter J., Kremizki C., Layman D.,
Leonard S., Sun H., Fulton L., Nash W., Miner T., Minx P.,
Delehaunty K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
Spieth J., Wilson R.K.
RT "Comparison of genome degradation in Paratyphi A and Typhi, human-
restricted serovars of Salmonella enterica that cause typhoid.";
RL Nat. Genet. 36:1268-1274(2004).
DR EMBL; CP000026; AAV78544.1; -; Genomic_DNA.
DR SMR; Q5PF15; 4-329.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003684; F:damaged DNA binding; IEA.
DR GO; GO:0008094; F:DNA-dependent ATPase activity; IEA.
DR GO; GO:0017111; F:nucleoside-triphosphatase activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PR00142; RECA.
DR ProDom; PD000229; RecA; 1.
DR SMART; SMO0382; AAA; 1.
DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS0162; RECA_2; 1.
DR PROSITE; PS0163; RECA_3; 1.
KW Complete proteome.
SQ SEQUENCE 353 AA; 37962 MW; AD15A0C48E7CF6FD CRC64;

Query Match 97.7%; Score 1639; DB 2; Length 353;
Best Local Similarity 97.0%; Pred. No. 2.2e-101;
Matches 325; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

```
Qy 1 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSMVDVETISTGSLSDIALGAGGLPMGR 60
Db 1 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSMVDVETISTGSLSDIALGAGGLPMGR 61
Qy 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 121
Qy 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEGSHMGLAARMMSQAMRKLAG 180
Db 122 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEGSHMGLAARMMSQAMRKLAG 181
Qy 181 NLKQSNLTLLIFNQIRMKIGWFGNPEPTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240
Db 182 NLKQSNLTLLIFNQIRMKIGWFGNPEPTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 241
Qy 241 ETRVVKVKNKIAAPFKQAEFQILYEGEGINFYGGELVDLGKVKLEIEKAGAWSYKGEKIGQ 300
Db 242 ETRVVKVKNKIAAPFKQAEFQILYEGEGINFYGGELVDLGKVKLEIEKAGAWSYNGEKIGQ 301
Qy 301 GKANATAWLKONPETAKIEKKVRELLLSNPNSTP 335
Db 302 GKANATWTKENPATATAKEIEKRVRELLLSNQATP 336

RESULT 10
RECA_PHOLL
ID RECA PHOLL STANDARD; PRT; 355 AA.
AC Q7N7A6;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocuNames=plu1249;
OS Photorhabdus luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
OX NCBI_TaxID=141679;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=22957627; PubMed=14528314; DOI=10.1038/nbt886;
RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,
RA Taurit S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-F.,
RA Dassa E., Derose R., Derzelle S., Freyssinet G., Gaudriault S.,
RA Medigue C., Lanois A., Powell K., Siguer P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.,
RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
RT luminescens."
RL Nat. Biotechnol. 21:1307-1313(2003).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; BX571863; CAE13543.1; -; Genomic_DNA.
DR SMR; Q7N7A6; 4-329.
DR PhotoList; plu1249; -.
DR HAMAP; MF_002688; -; 1.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PR00142; RecA.
DR ProDom; PD000229; RecA; 1.
```

```
DR TIGRFAMS; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS50162; RECA_2; 1.
DR PROSITE; PS50163; RECA_3; 1.
KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
FT NP_BIND 67 74 ATP (By similarity).
SQ SEQUENCE 355 AA; 38381 MW; PD659ACED827AE9 CRC64;

Query Match 92.2%; Score 1547; DB 1; Length 355;
Beat Local Similarity 92.1%; Pred. No. 3.1e-95;
Matches 304; Conservative 15; Mismatches 11; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSMVDVETISTGSLSDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSMVDVETISTGSLSDIALGAGGLPMGR 61
Qy 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 121
Qy 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEGSHMGLAARMMSQAMRKLAG 180
Db 122 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEGSHMGLAARMMSQAMRKLAG 181
Qy 181 NLKQSNLTLLIFNQIRMKIGWFGNPEPTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240
Db 182 NLKQSNLTLLIFNQIRMKIGWFGNPEPTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 241
Qy 241 ETRVVKVKNKIAAPFKQAEFQILYEGEGINFYGGELVDLGKVKLEIEKAGAWSYKGEKIGQ 300
Db 242 ETRVVKVKNKIAAPFKQAEFQILYEGEGINFYGGELVDLGKVKLEIEKAGAWSYNGEKIGQ 301
Qy 301 GKANATAWLKONPETAKIEKKVRELLLSN 330
Db 302 GKANATWTKENPATATAKEIEKRVRELLLSN 331

RESULT 11
RECA_ERWCT
ID RECA_ERWCT STANDARD; PRT; 357 AA.
AC Q6D1S8;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocuNames=ECA3369;
OS Erwinia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=SCRI 1043 / ATCC BAA-672;
RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebahia M., Pritchard L., Holden M.T.G., Hyman L.J.,
RA Holava M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,
RA Atkin K., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
RT "Genome sequence of the enterobacterial phytopathogen Erwinia
RT carotovora subsp. atroseptica and characterization of virulence
RT factors."
RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
```


CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

DR EMBL; BX950851; CAG76267.1; -; Genomic_DNA.

DR SMR; Q6D188; 4-329.

DR HAMAP; MF_00268; -; 1.

DR InterPro; IPR003593; AAA ATPase.

DR InterPro; IPR001553; RecA.

DR Pfam; PF00154; RecA; 1.

DR PRINTS; PD00142; RecA.

DR ProDom; PD000229; RecA; 1.

DR SMART; SM00382; AAA; 1.

DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.

DR PROSITE; PS00321; RecA_1; 1.

DR PROSITE; PS00162; RecA_2; 1.

DR PROSITE; PS00163; RecA_3; 1.

KW ATP-binding; Complete proteome; DNA damage; DNA recombination;

KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.

FT NP BIND 67 74 ATP (By similarity).

SQ SEQUENCE 357 AA; 38284 MW; 48F1FFB8232E0098 CRC64;

Query Match 91.6%; Score 1537; DB 1; Length 357;

Best Local Similarity 91.6%; Pred. No. 1.4e-94;

Matches 304; Conservative 14; Mismatches 14; Indels 0; Gaps 0;

QY 1 AIDENKQALAAALGOIEKQFGKSGIMRLGDRSDMDVTITGSLSLDIALGAGLPNGR 60

Db 2 AIDENKQALAAALGOIEKQFGKSGIMRLGDRSDMDVTITGSLSLDIALGAGLPNGR 61

QY 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCFIDAEHALDPIYAKLGVDIDNLLCSQPD 120

Db 62 IVEIYGPSSGKTTLTQVIAAAQREGKTCFIDAEHALDPIYAKLGVDIDNLLCSQPD 121

QY 121 TGEQALEICDALARSGVDVIVDSVAALTPKAEIEGIGSHGLAARMSQAMRKLKAG 180

Db 122 TGEQALEICDALARSGVDVIVDSVAALTPKAEIEGIGSHGLAARMSQAMRKLKAG 181

QY 181 NLKQSNLTLLIPIQIRMKIGVMFGNPTTTCGNALKFYASVRLDIRRIGAVKEGVVGS 240

Db 182 NLKQSNLTLLIPIQIRMKIGVMFGNPTTTCGNALKFYASVRLDIRRIGAVKEGVVGS 241

QY 241 ETRVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGKVKLEKAGAWTSYKGEIKIGQ 300

Db 242 ETRVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGKVKLEKAGAWTSYKGEIKIGQ 301

QY 301 GKANATATLNDKNDPTAKEIEKKVRELLISNPN 332

Db 302 GKANACNFKLNPTISAELDKKREMLLHKGN 333

RESULT 12

RECA YERPE

ID RECA YERPE

AC P37858;

DT 01-OCT-1994 (Rel. 30, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE RecA protein (Recombinase A).

GN Name=recA; OrderedLocNames=YPO3307, Y0881, YP0379;

OS Versinia pestis.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Versinia.

OX NCBI_TaxID=632;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=231;

RA Kryukov V.M., Suchkov I.Y., Sazykin I.S., Mishankin B.N.;

RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.

RN [2]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=CO-92 / Biovar Orientalis;

RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;

RA Parhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,

RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,

RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Rarraga A.-M.,

RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,

RA Fellwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,

RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,

RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;

RT "Genome sequence of Versinia pestis, the causative agent of plague.";

RL Nature 413:523-527(2001).

RN [3]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=KIM5 / Biovar Mediaevalis;

RX MEDLINE=22137863; PubMed=12142430;

RX DOI=10.1128/JB.184.16.4601-4611.2002;

RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,

RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,

RA Fetherston J.D., Lindler L.E., Zhubaker R.R., Plano G.V.,

RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,

RA Perry R.D.;

RT "Genome sequence of Versinia pestis KIM.";

RL J. Bacteriol. 184:4601-4611(2002).

RN [4]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=91001 / Biovar Mediaevalis;

RX PubMed=15368893;

RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,

RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,

RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,

RA Yang H., Wang J., Huang P., Yang R.;

RT "Complete genome sequence of Versinia pestis strain 91001, an isolate

avirulent to humans.";

RL DNA Res. 11:179-197(2004).

CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of

single-stranded DNA, the ATP-dependent uptake of single-stranded

DNA by duplex DNA, and the ATP-dependent hybridization of

homologous single-stranded DNAs. It interacts with LexA causing

its activation and leading to its autocatalytic cleavage.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -!- SIMILARITY: Belongs to the recA family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use as long as its content is in no way modified and this statement is not

removed.

CC EMBL; X75336; CAA53084.1; -; Genomic DNA.

DR EMBL; AJ414156; CAC92539.1; -; Genomic DNA.

DR EMBL; AE013691; AAM84465.1; -; Genomic DNA.

DR EMBL; AE017128; AAS60652.1; -; Genomic DNA.

DR PIR; AG0401; AG0401.

DR PIR; S37586; S37586.

DR HSPP; P03017; 2REB.

DR SMR; P37858; 3-328.

DR HAMAP; MF_00268; -; 1.

DR InterPro; IPR003593; AAA ATPase.

DR InterPro; IPR001553; RecA.

DR Pfam; PF00154; RecA; 1.

DR PRINTS; PR00142; RecA.

DR ProDom; PD000229; RecA; 1.

DR SMART; SM00382; AAA; 1.

DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.

DR PROSITE; PS00321; RecA_1; 1.

DR PROSITE; PS00162; RecA_2; 1.

DR PROSITE; PS00163; RecA_3; 1.

KW ATP-binding; Complete proteome; DNA damage; DNA recombination;

KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.

FT INIT MET 0 0 By similarity.

FT NP BIND 66 73 ATP (By similarity).

FT CONFLICT 281 281 H -> L (in Ref. 1).

FT CONFLICT 294 294 N -> Y (in Ref. 1).
 FT CONFLICT 315 315 I -> N (in Ref. 1).
 SQ SEQUENCE 355 AA; 37755 MW; F800D2D1AD32AB81 CRC64;
 Query Match 91.4%; Score 1534; DB 1; Length 355;
 Best Local Similarity 91.9%; Pred. No. 2.3e-94;
 Matches 305; Conservative 14; Mismatches 13; Indels 0; Gaps 0;
 QY 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSMVETISTGSLSLDIALGAGGLPMGR 60
 DB 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSMVETISTGSLSLDIALGAGGLPMGR 60
 QY 61 IVEIYGPSSSGKTTLTQVIAAAQREGKTCFADAEHALDPIYAKLGVDIDNLLCSQPD 120
 DB 61 IVEIYGPSSSGKTTLTQVIAAAQREGKTCFADAEHALDPIYAKLGVDIDNLLCSQPD 120
 QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEGISHMGLAARMMSQAMRKLKAG 180
 DB 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEGISHMGLAARMMSQAMRKLKAG 180
 QY 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
 DB 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
 QY 241 ETRVVKVKNKIAAPFKQAEFOILYEGEINFGYELVDLGKVKELIEKAGAWYSYKGEKIGQ 300
 DB 241 ETRVVKVKNKIAAPFKQAEFOILYEGEINFGYELVDLGKVKELIEKAGAWYSYKGEKIGQ 300
 QY 301 GKANATWLKONPETAKEIEKKVRELLLSNP 332
 DB 301 GKANASNYLKENPAIAELDKKREMLLNGN 332

RESULT 13

RECA YERPS STANDARD; PRT; 356 AA.
 ID RECA YERPS STANDARD; PRT; 356 AA.
 AC Q66E70;
 DT 10-MAY-2005 (Rel. 47, Created)
 DT 10-MAY-2005 (Rel. 47, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE RecA protein (Recombinase A).
 GN Name=recA; OrderedLocusName=YPTB0823;
 OS Yersinia pseudotuberculosis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=633;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=IP32953 / Serotype I;
 RX PubMed=15358858; DOI=10.1073/pnas.040401201;
 RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
 RA Regala W.M., Georgescu A.M., Verges L.M., Land M.L., Motin V.L.,
 RA Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
 RA Simonet M., Chenal-Francoise V., Souza B., Dacheux D., Elliott J.M.,
 RA Derbise A., Hauser L.J., Garcia E.;
 RT "Insights into the evolution of Yersinia pestis through whole-genome
 comparison with Yersinia pseudotuberculosis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
 CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
 CC single-stranded DNA, and the ATP-dependent uptake of single-stranded
 CC DNA by duplex DNA, and the ATP-dependent hybridization of
 CC homologous single-stranded DNAs. It interacts with LexA causing
 CC its activation and leading to its autocatalytic cleavage (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the RecA family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

DR EMBL; BX936398; CAH20063.1; -; Genomic_DNA.
 DR SMR; Q66E70; 4-329.
 DR HAMAP; MF_00268; -; 1.
 DR InterPro; IPR003593; AAA_ATPase.
 DR Pfam; PF001553; RecA.
 DR Pfam; PF00154; RecA; 1.
 DR PRINTS; PD00142; RECA.
 DR ProDom; PD000229; RECA; 1.
 DR SMART; SM00382; AAA; 1.
 DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
 DR PROSITE; PS00321; RECA_1; 1.
 DR PROSITE; PS0162; RECA_2; 1.
 DR PROSITE; PS0163; RECA_3; 1.
 KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
 KW ATP-repair; DNA-binding; Nucleotide-binding; SOS response.
 FT NP BIND 67 74 ATP (By similarity).
 SQ SEQUENCE 356 AA; 37914 MW; DIC0D90F58C642B CRC64;
 Query Match 91.4%; Score 1534; DB 1; Length 356;
 Best Local Similarity 91.9%; Pred. No. 2.3e-94;
 Matches 305; Conservative 14; Mismatches 13; Indels 0; Gaps 0;
 QY 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSMVETISTGSLSLDIALGAGGLPMGR 60
 DB 2 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSMVETISTGSLSLDIALGAGGLPMGR 61
 QY 61 IVEIYGPSSSGKTTLTQVIAAAQREGKTCFADAEHALDPIYAKLGVDIDNLLCSQPD 120
 DB 62 IVEIYGPSSSGKTTLTQVIAAAQREGKTCFADAEHALDPIYAKLGVDIDNLLCSQPD 121
 QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEGISHMGLAARMMSQAMRKLKAG 180
 DB 122 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEGISHMGLAARMMSQAMRKLKAG 181
 QY 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
 DB 182 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 241
 QY 241 ETRVVKVKNKIAAPFKQAEFOILYEGEINFGYELVDLGKVKELIEKAGAWYSYKGEKIGQ 300
 DB 242 ETRVVKVKNKIAAPFKQAEFOILYEGEINFGYELVDLGKVKELIEKAGAWYSYKGEKIGQ 301
 QY 301 GKANATWLKONPETAKEIEKKVRELLLSNP 332
 DB 302 GKANASNYLKENPAIAELDKKREMLLNGN 333
 RESULT 14
 RECA ENTAG STANDARD; PRT; 353 AA.
 ID RECA ENTAG STANDARD; PRT; 353 AA.
 AC P33037;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE RecA protein (Recombinase A).
 GN Name=recA;
 OS Enterobacter agglomerans (Erwinia herbicola) (Pantoea agglomerans).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Pantoea.
 OX NCBI_TaxID=549;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
 RA Rappold C.S.J., Klingmueller W.;
 RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
 CC single-stranded DNA, the ATP-dependent uptake of single-stranded
 CC DNA by duplex DNA, and the ATP-dependent hybridization of
 CC homologous single-stranded DNAs. It interacts with LexA causing
 CC its activation and leading to its autocatalytic cleavage.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the RecA family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

DR EMBL; L03291; AAA91766.1; -; Genomic_DNA.
 DR PIR; S1481; S31481.
 DR HSSP; P03017; 2REB.
 DR SMR; P33037; 3-328.
 DR HAMAP; MF_00268; -; 1.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR001553; RecA.
 DR Pfam; PF00154; RecA; 1.
 DR PRINTS; PR00142; RECA.
 DR ProDom; PD000229; RecA; 1.
 DR SMART; SM00382; AAA; 1.
 DR TIGRFAMs; TIGR02012; tigrfam_reca; 1.
 DR PROSITE; PS00321; RECA_1; 1.
 DR PROSITE; PS50162; RECA_2; 1.
 DR PROSITE; PS50163; RECA_3; 1.
 KW ATP-binding; DNA damage; DNA recombination; DNA repair; DNA-binding;
 KW Nucleotide-binding; SOS response.
 FT INIT MET 0 By similarity.
 FT NP_BIND 66 73 ATP (By similarity).
 SQ SEQUENCE 353 AA; 37767 MW; 960F1F2698CAC46B CRC64;

Query Match 91.3%; Score 1532; DB 1; Length 353;
 Best Local Similarity 92.1%; Pred. No. 3.1e-94;
 Matches 303; Conservative 15; Mismatches 11; Indels 0; Gaps 0;
 Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVETISTGSLDIALGAGGLPMGR 60
 Db 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVETISTGSLDIALGAGGLPMGR 60
 Qy 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPYARKLGVDIDNLLCSQPD 120
 Db 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPYARKLGVDIDNLLCSQPD 120
 Qy 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGEGIDSHMGLAARMMSQAMRKLKAG 180
 Db 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGEGIDSHMGLAARMMSQAMRKLKAG 180
 Qy 181 NLKQSNLTLLIFINQIRMKIGVMFNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVGVS 240
 Db 181 NLKQSNLTLLIFINQIRMKIGVMFNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVGVS 240
 Qy 241 ETRVKVKNKTAAPPKQAEFOILYGEINFGELVDLGKVEKLEKAGAWYSYNGEKTIGQ 300
 Db 241 ETRVKVKNKTAAPPKQAEFOILYGEINFGELVDLGKVEKLEKAGAWYSYNGEKTIGQ 300
 Qy 301 GKANATAWLKDNPETAKEIEKKVRELLLS 329
 Db 301 GKANSCNYLKENPKVAELDKKLRDMLLS 329

RESULT 15

RECA_SERMA STANDARD; PRT; 353 AA.
 AC P17479;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE ReCA protein (Recombinase A).
 GN Name=reca;
 OS Serratia marcescens.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Serratia.
 OC NCBI_TaxID=615;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
 RC STRAIN=SM6;
 RX MEDLINE=90094239; PubMed=2152908;
 RA Ball T.K., Wasmuth C.R., Braunagel S.C., Benedik M.J.;

RT "Expression of Serratia marcescens extracellular proteins requires
 recA".
 RL J. Bacteriol. 172:342-349(1990).
 CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
 CC single-stranded DNA, the ATP-dependent uptake of single-stranded
 CC DNA by duplex DNA, and the ATP-dependent hybridization of
 CC homologous single-stranded DNAs. It interacts with lexA causing
 CC its activation and leading to its autocatalytic cleavage.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the recA family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

CC EMBL; M22935; AAA26567.1; -; Genomic_DNA.
 CC HSSP; P03017; 2REB.
 CC SMR; P17479; 3-328.
 CC HAMAP; MF_00268; -; 1.
 CC InterPro; IPR003593; AAA_ATPase.
 CC InterPro; IPR001553; RecA.
 CC Pfam; PF00154; RecA; 1.
 CC PRINTS; PR00142; RECA.
 CC ProDom; PD000229; RecA; 1.
 CC SMART; SM00382; AAA; 1.
 CC TIGRFAMs; TIGR02012; tigrfam_reca; 1.
 CC PROSITE; PS00321; RECA_1; 1.
 CC PROSITE; PS50162; RECA_2; 1.
 CC PROSITE; PS50163; RECA_3; 1.
 KW ATP-binding; DNA damage; DNA recombination; DNA repair; DNA-binding;
 KW Nucleotide-binding; SOS response.
 FT INIT MET 0 By similarity.
 FT NP_BIND 66 73 ATP (By similarity).
 SQ SEQUENCE 353 AA; 37778 MW; A97369360970F814 CRC64;

Query Match 91.0%; Score 1527; DB 1; Length 353;
 Best Local Similarity 92.4%; Pred. No. 6.6e-94;
 Matches 303; Conservative 13; Mismatches 12; Indels 0; Gaps 0;
 Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVETISTGSLDIALGAGGLPMGR 60
 Db 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVETISTGSLDIALGAGGLPMGR 60
 Qy 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPYARKLGVDIDNLLCSQPD 120
 Db 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPYARKLGVDIDNLLCSQPD 120
 Qy 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGEGIDSHMGLAARMMSQAMRKLKAG 180
 Db 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGEGIDSHMGLAARMMSQAMRKLKAG 180
 Qy 181 NLKQSNLTLLIFINQIRMKIGVMFNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVGVS 240
 Db 181 NLKQSNLTLLIFINQIRMKIGVMFNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVGVS 240
 Qy 241 ETRVKVKNKTAAPPKQAEFOILYGEINFGELVDLGKVEKLEKAGAWYSYNGEKTIGQ 300
 Db 241 ETRVKVKNKTAAPPKQAEFOILYGEINFGELVDLGKVEKLEKAGAWYSYNGEKTIGQ 300
 Qy 301 GKANATAWLKDNPETAKEIEKKVRELLLS 328
 Db 301 GKANACNFKENPAIAELDKKLRDMLLS 328

Search completed: February 16, 2006, 01:24:00
 Job time : 153 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 01:24:19 ; Search time 32.5 Seconds
(without alignments)
852.195 Million cell updates/sec

Title: US-10-733-782-1
Perfect score: 1678
Sequence: 1 AIDENKQKALAAALGQIEKQ.....AKIEKKVRELLSNPNSTP 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5 COMB.pap:*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pap:*
3: /cgn2_6/ptodata/1/iaa/H COMB.pap:*
4: /cgn2_6/ptodata/1/iaa/PCUS COMB.pap:*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pap:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1678	100.0	353	1	US-08-097-831-6
2	1678	100.0	358	2	US-09-626-410-7
3	1678	100.0	358	2	US-09-626-410-9
4	1678	100.0	358	2	US-09-626-410-14
5	1678	100.0	358	2	US-09-116-188-7
6	1678	100.0	358	2	US-09-116-188-9
7	1678	100.0	358	2	US-09-116-188-14
8	1678	100.0	358	2	US-09-626-047-9
9	1678	100.0	358	2	US-09-626-047-14
10	1678	100.0	358	2	US-09-626-047-9
11	1678	100.0	358	2	US-09-626-047-14
12	1678	100.0	358	2	US-09-626-343-7
13	1678	100.0	358	2	US-09-626-343-9
14	1678	100.0	358	2	US-09-354-922-8
15	1678	100.0	358	2	US-09-354-922-10
16	1678	100.0	358	2	US-09-354-922-15
17	1678	100.0	358	2	US-09-516-051-7
18	1678	100.0	358	2	US-09-516-051-9
19	1678	100.0	358	2	US-09-516-051-14
20	1678	100.0	358	2	US-09-516-695B-8
21	1678	100.0	358	2	US-09-516-695B-10
22	1678	100.0	358	2	US-09-516-695B-15
23	1673	99.7	358	2	US-09-626-410-8
24	1673	99.7	358	2	US-09-626-410-11
25	1673	99.7	358	2	US-09-116-188-8
26	1673	99.7	358	2	US-09-116-188-11
27	1673	99.7	358	2	US-09-626-047-8

28	1673	99.7	358	2	US-09-626-047-11	Sequence 11, Appl
29	1673	99.7	358	2	US-09-626-343-8	Sequence 8, Appl
30	1673	99.7	358	2	US-09-626-343-11	Sequence 11, Appl
31	1673	99.7	358	2	US-09-354-922-9	Sequence 9, Appl
32	1673	99.7	358	2	US-09-354-922-12	Sequence 12, Appl
33	1673	99.7	358	2	US-09-516-051-8	Sequence 8, Appl
34	1673	99.7	358	2	US-09-516-051-11	Sequence 11, Appl
35	1673	99.7	358	2	US-09-516-695B-9	Sequence 9, Appl
36	1673	99.7	358	2	US-09-516-695B-12	Sequence 12, Appl
37	1660	98.9	358	2	US-09-626-410-12	Sequence 12, Appl
38	1660	98.9	358	2	US-09-116-188-12	Sequence 12, Appl
39	1660	98.9	358	2	US-09-626-047-12	Sequence 12, Appl
40	1660	98.9	358	2	US-09-626-343-12	Sequence 12, Appl
41	1660	98.9	358	2	US-09-354-922-13	Sequence 13, Appl
42	1660	98.9	358	2	US-09-516-051-12	Sequence 12, Appl
43	1660	98.9	358	2	US-09-516-695B-13	Sequence 13, Appl
44	1659	98.9	358	2	US-09-626-410-10	Sequence 10, Appl
45	1659	98.9	358	2	US-09-116-188-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-097-831-6
; Sequence 6, Application US/08097831
; Patent No. 5510473
; GENERAL INFORMATION:
; APPLICANT: Camerini-Otero, Rafael D.
; TITLE OF INVENTION: Cloning and Expression of Taq reca
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/097,831
; APPLICATION NUMBER: US/08/097,831
; FILING DATE: 19930726
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E.
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: NIH066.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
; TELEFAX: 714-760-9502
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-097-831-6

Query Match 100.0%; Score 1678; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 2.5e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQKGMRLCEDSRMDVETISTGSLSLDIAGAGLPMGR 60
|||||

Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 61
Qy 61 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 121
Qy 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 181
Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
Db 182 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 241
Qy 241 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVEKLEIKAGAWYSYKGEKIGQ 300
Db 242 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVEKLEIKAGAWYSYKGEKIGQ 301
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 302 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 336

RESULT 2
US-09-626-410-7
; Sequence 7, Application US/09626410
; Patent No. 6287862
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/626,410
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 7
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-626-410-7

Query Match 100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 126
Qy 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 186
Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
Db 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 246

Qy 241 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVEKLEIKAGAWYSYKGEKIGQ 300
Db 247 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVEKLEIKAGAWYSYKGEKIGQ 306
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 3
US-09-626-410-9
; Sequence 9, Application US/09626410
; Patent No. 6287862
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/626,410
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 9
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-626-410-9

Query Match 100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 126
Qy 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 186
Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
Db 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 246
Qy 241 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVEKLEIKAGAWYSYKGEKIGQ 300
Db 247 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVEKLEIKAGAWYSYKGEKIGQ 306
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 4
US-09-626-410-14
; Sequence 14, Application US/09626410
; Patent No. 6287862
; GENERAL INFORMATION:

APPLICANT: DEL CARDAYRE, STEPHEN
APPLICANT: TOBIN, MATTHEW
APPLICANT: STEMMER, WILLEM P. C.
APPLICANT: NESS, JON E.
APPLICANT: MINSHULL, JEREMY
APPLICANT: PATTEN, PHILLIP
APPLICANT: SUBRAMANIA, VENKITSWARAN
APPLICANT: CASTLE, LINDA
APPLICANT: KREBER, CLAUD M.
APPLICANT: BASS, STEVE
TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
FILE REFERENCE: 02-020720US
CURRENT APPLICATION NUMBER: US/09/626,410
CURRENT FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 09/116,188
PRIOR FILING DATE: 07-15-1998
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 358
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: consensus
OTHER INFORMATION: e. coli sequence
US-09-626-410-14

Query Match 100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPSSGKTTLTLOVIAAAREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPSSGKTTLTLOVIAAAREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 126
Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 186
Qy 181 NLKQSTLLIFINQIRMKIGVFGNPETTTGNNALKFYASVRLDIRRIGAVKEGENNVGS 240
Db 187 NLKQSTLLIFINQIRMKIGVFGNPETTTGNNALKFYASVRLDIRRIGAVKEGENNVGS 246
Qy 241 ETRVVKVKNKTAAPFKQAEFQILYGEINFGELVDLGKVKLEKAGAWSYKGEKIGQ 300
Db 247 ETRVVKVKNKTAAPFKQAEFQILYGEINFGELVDLGKVKLEKAGAWSYKGEKIGQ 306
Qy 301 GKANATLWLNKDPETAKEIEKKVRELLLSNPNSTP 335
Db 307 GKANATLWLNKDPETAKEIEKKVRELLLSNPNSTP 341

RESULT 5
US-09-116-188-7
Sequence 7, Application US/09116188
Patent No. 6326204
GENERAL INFORMATION:
APPLICANT: DEL CARDAYRE, STEPHEN
APPLICANT: TOBIN, MATTHEW
APPLICANT: STEMMER, WILLEM P. C.
APPLICANT: NESS, JON E.
APPLICANT: MINSHULL, JEREMY
APPLICANT: PATTEN, PHILLIP
APPLICANT: SUBRAMANIA, VENKITSWARAN
APPLICANT: CASTLE, LINDA
APPLICANT: KREBER, CLAUD M.
APPLICANT: BASS, STEVE
TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE

TITLE OF INVENTION: SEQUENCE RECOMBINATION
FILE REFERENCE: 02-020720US
CURRENT APPLICATION NUMBER: US/09/116,188
CURRENT FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 358
TYPE: PRT
ORGANISM: Escherichia coli
US-09-116-188-7
Query Match 100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPSSGKTTLTLOVIAAAREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPSSGKTTLTLOVIAAAREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 126
Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 186
Qy 181 NLKQSTLLIFINQIRMKIGVFGNPETTTGNNALKFYASVRLDIRRIGAVKEGENNVGS 240
Db 187 NLKQSTLLIFINQIRMKIGVFGNPETTTGNNALKFYASVRLDIRRIGAVKEGENNVGS 246
Qy 241 ETRVVKVKNKTAAPFKQAEFQILYGEINFGELVDLGKVKLEKAGAWSYKGEKIGQ 300
Db 247 ETRVVKVKNKTAAPFKQAEFQILYGEINFGELVDLGKVKLEKAGAWSYKGEKIGQ 306
Qy 301 GKANATLWLNKDPETAKEIEKKVRELLLSNPNSTP 335
Db 307 GKANATLWLNKDPETAKEIEKKVRELLLSNPNSTP 341

RESULT 6
US-09-116-188-9
Sequence 9, Application US/09116188
Patent No. 6326204
GENERAL INFORMATION:
APPLICANT: DEL CARDAYRE, STEPHEN
APPLICANT: TOBIN, MATTHEW
APPLICANT: STEMMER, WILLEM P. C.
APPLICANT: NESS, JON E.
APPLICANT: MINSHULL, JEREMY
APPLICANT: PATTEN, PHILLIP
APPLICANT: SUBRAMANIA, VENKITSWARAN
APPLICANT: CASTLE, LINDA
APPLICANT: KREBER, CLAUD M.
APPLICANT: BASS, STEVE
TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
FILE REFERENCE: 02-020720US
CURRENT APPLICATION NUMBER: US/09/116,188
CURRENT FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 358
TYPE: PRT
ORGANISM: Escherichia coli
US-09-116-188-9

Query Match 100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 66
QY 61 IVEIYGPESSGKTTTLTQVIAAQAQREGKTCAPIDAHAHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPESSGKTTTLTQVIAAQAQREGKTCAPIDAHAHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSQAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARSQAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 186
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
QY 241 ETRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKLEIKAGAWYSYKGEKIGQ 300
DB 247 ETRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKLEIKAGAWYSYKGEKIGQ 306
QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
DB 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 7
US-09-116-188-14
; Sequence 14, Application US/09116188
; Patent No. 6326204
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBBER, CLAUD M.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: US/09/116,188
; CURRENT FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; OTHER INFORMATION: e. coli sequence
US-09-116-188-14

Query Match 100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 66
QY 61 IVEIYGPESSGKTTTLTQVIAAQAQREGKTCAPIDAHAHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPESSGKTTTLTQVIAAQAQREGKTCAPIDAHAHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSQAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARSQAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 186
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240

DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
QY 241 ETRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKLEIKAGAWYSYKGEKIGQ 300
DB 247 ETRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKLEIKAGAWYSYKGEKIGQ 306
QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
DB 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 8
US-09-626-047-7
; Sequence 7, Application US/09626047
; Patent No. 6335198
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBBER, CLAUD M.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/626,047
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-626-047-7
Query Match 100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 66
QY 61 IVEIYGPESSGKTTTLTQVIAAQAQREGKTCAPIDAHAHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPESSGKTTTLTQVIAAQAQREGKTCAPIDAHAHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSQAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARSQAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 186
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
QY 241 ETRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKLEIKAGAWYSYKGEKIGQ 300
DB 247 ETRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKLEIKAGAWYSYKGEKIGQ 306
QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
DB 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 9
US-09-626-047-9

; Sequence 9, Application US/09626047
; Patent No. 6335198
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: SEQUENCE RECOMBINATION
; CURRENT APPLICATION NUMBER: US/09/626,047
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-626-047-9

Query Match 100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 66
QY 61 IVEIYGPESSGKTTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPESSGKTTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARGAVDVIIVDSVAALTPKAEIEGEIGDISHMGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARGAVDVIIVDSVAALTPKAEIEGEIGDISHMGLAARMMSQAMRKLKAG 186
QY 181 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
DB 187 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 246
QY 241 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKELIEKAGAWSYKGEKIQ 300
DB 247 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKELIEKAGAWSYKGEKIQ 306
QY 301 GKANATAMKONPETAKEIEKKVRELLLSNPSTP 335
DB 307 GKANATAMKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 10
US-09-626-047-14
; Sequence 14, Application US/09626047
; Patent No. 6335198
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE

; TITLE OF INVENTION: SEQUENCE RECOMBINATION
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/626,047
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; OTHER INFORMATION: e. coli sequence
US-09-626-047-14

Query Match 100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 66
QY 61 IVEIYGPESSGKTTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPESSGKTTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARGAVDVIIVDSVAALTPKAEIEGEIGDISHMGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARGAVDVIIVDSVAALTPKAEIEGEIGDISHMGLAARMMSQAMRKLKAG 186
QY 181 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
DB 187 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 246
QY 241 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKELIEKAGAWSYKGEKIQ 300
DB 247 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKELIEKAGAWSYKGEKIQ 306
QY 301 GKANATAMKONPETAKEIEKKVRELLLSNPSTP 335
DB 307 GKANATAMKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 11
US-09-626-343-7
; Sequence 7, Application US/09626343
; Patent No. 6352859
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/626,343
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 358
; TYPE: PRT

Db 247 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLEKAGAWSYKGEKIGQ 306
Qy 301 GKANATAWKONPETAKEIEKKVRELLSNPNSTP 335
Db 307 GKANATAWKONPETAKEIEKKVRELLSNPNSTP 341

RESULT 14

US-09-354-922-8

; Sequence 8, Application US/09354922

; Patent No. 6379964

; GENERAL INFORMATION:

; APPLICANT: DEL CARDAYRE, STEPHEN

; APPLICANT: TOBIN, MATTHEW

; APPLICANT: STEMMER, WILLEM P.C.

; APPLICANT: NESS, JON E.

; APPLICANT: MINSHULL, JEREMY

; APPLICANT: PATTEN, PHILLIP

; APPLICANT: SUBRAMANIAN, VENKITESWATAN

; APPLICANT: CASTLE, LINDA A.

; APPLICANT: KREBER, CLAUD M.

; APPLICANT: BASS, STEVE

; APPLICANT: ZHANG, YING-XIN

; APPLICANT: COX, TONY

; APPLICANT: HUISMAN, GJALT

; APPLICANT: AFFHOLTER, JOSEPH A.

; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE

; FILE REFERENCE: 02-020730US

; CURRENT APPLICATION NUMBER: US/09/354,922

; CURRENT FILING DATE: 1999-07-15

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 8

; LENGTH: 358

; TYPE: PRT

; ORGANISM: Escherichia coli

US-09-354-922-8

Query Match 100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGOIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPWGR 60
Db 7 AIDENKOKALAAALGOIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPWGR 66

Qy 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVLDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVLDNLLCSQPD 126

Qy 121 TGEQALEICDALARSGADVIVDVSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGADVIVDVSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 186

Qy 181 NLKOSNTLLIFINQIRMKIGVMFGNPETTTGCGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKOSNTLLIFINQIRMKIGVMFGNPETTTGCGNALKFYASVRLDIRRIGAVKEGENVVGS 246

Qy 241 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLEKAGAWSYKGEKIGQ 300
Db 247 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLEKAGAWSYKGEKIGQ 306

Qy 301 GKANATAWKONPETAKEIEKKVRELLSNPNSTP 335
Db 307 GKANATAWKONPETAKEIEKKVRELLSNPNSTP 341

RESULT 15

US-09-354-922-10

; Sequence 10, Application US/09354922

; Patent No. 6379964

; GENERAL INFORMATION:

; APPLICANT: DEL CARDAYRE, STEPHEN

; APPLICANT: TOBIN, MATTHEW

; APPLICANT: STEMMER, WILLEM P.C.

; APPLICANT: NESS, JON E.

; APPLICANT: MINSHULL, JEREMY

; APPLICANT: PATTEN, PHILLIP

; APPLICANT: SUBRAMANIAN, VENKITESWATAN

; APPLICANT: CASTLE, LINDA A.

; APPLICANT: KREBER, CLAUD M.

; APPLICANT: BASS, STEVE

; APPLICANT: ZHANG, YING-XIN

; APPLICANT: COX, TONY

; APPLICANT: HUISMAN, GJALT

; APPLICANT: YUAN, LING

; APPLICANT: AFFHOLTER, JOSEPH A.

; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE

; FILE REFERENCE: 02-020730US

; CURRENT APPLICATION NUMBER: US/09/354,922

; CURRENT FILING DATE: 1999-07-15

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 10

; LENGTH: 358

; TYPE: PRT

; ORGANISM: Escherichia coli

US-09-354-922-10

Query Match

100.0%; Score 1678; DB 2; Length 358;

Best Local Similarity 100.0%; Pred. No. 2.6e-167;

Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGOIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPWGR 60
Db 7 AIDENKOKALAAALGOIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPWGR 66

Qy 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVLDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVLDNLLCSQPD 126

Qy 121 TGEQALEICDALARSGADVIVDVSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGADVIVDVSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 186

Qy 181 NLKOSNTLLIFINQIRMKIGVMFGNPETTTGCGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKOSNTLLIFINQIRMKIGVMFGNPETTTGCGNALKFYASVRLDIRRIGAVKEGENVVGS 246

Qy 241 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLEKAGAWSYKGEKIGQ 300
Db 247 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLEKAGAWSYKGEKIGQ 306

Qy 301 GKANATAWKONPETAKEIEKKVRELLSNPNSTP 335
Db 307 GKANATAWKONPETAKEIEKKVRELLSNPNSTP 341

Search completed: February 16, 2006, 01:26:08

Job time : 33.5 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 01:39:50 ; Search time 123 Seconds
(without alignments)
1137.990 Million cell updates/sec

Title: US-10-733-782-1
Perfect score: 1678
Sequence: 1 AIDENKQKALAAALGQIEKQ.....AKEIEKKVRELLSNPNSTP 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1678	100.0	325	4	US-10-733-782-1
2	1678	100.0	352	4	US-10-733-782-5
3	1678	100.0	353	4	US-10-282-122A-43319
4	1678	100.0	358	4	US-10-194-686-8
5	1678	100.0	358	4	US-10-194-686-10
6	1678	100.0	358	4	US-10-194-686-15
7	1674	99.8	325	4	US-10-733-782-3
8	1673	99.7	358	4	US-10-194-686-9
9	1673	99.7	358	4	US-10-194-686-12
10	1660	99.9	358	4	US-10-194-686-13
11	1659	98.9	358	4	US-10-194-686-11
12	1641	97.8	353	4	US-10-282-122A-75948
13	1613	96.1	352	4	US-10-282-122A-59525
14	1593	94.9	352	4	US-10-282-122A-55632
15	1534	91.4	356	4	US-10-282-122A-78524
16	1504	89.6	355	4	US-10-282-122A-68593
17	1467	87.4	351	4	US-10-282-122A-72895
18	1436	85.6	412	4	US-10-282-122A-77075
19	1353	80.6	354	4	US-10-282-122A-67415
20	1332	79.4	354	4	US-10-282-122A-58235
21	1268	75.6	346	4	US-10-282-122A-66505
22	1259	75.0	349	4	US-10-282-122A-44768
23	1257	74.9	355	4	US-10-282-122A-67942
24	1255	74.8	369	4	US-10-282-122A-69802
25	1245.5	74.2	349	4	US-10-282-122A-63392
26	1241.5	74.0	348	4	US-10-282-122A-61049
27	1239.5	73.9	342	4	US-10-751-928-20

RESULT 1

US-10-733-782-1
; Sequence 1, Application US/10733782
; Publication No. US20040157248A1
; GENERAL INFORMATION:
; APPLICANT: Cox, Michael
; APPLICANT: Lusetti, Shelley
; APPLICANT: Eggleter, Aimee
; TITLE OF INVENTION: RecA Mutants
; FILE REFERENCE: 960296.99501
; CURRENT APPLICATION NUMBER: US/10/733,782
; CURRENT FILING DATE: 2003-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 1
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-733-782-1

Query Match 100.0%; Score 1678; DB 4; Length 335;
Best Local Similarity 100.0%; Pred. No. 4.2e-152;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR	60
Db	1	AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR	60
Qy	61	I VEIYGPESSCKTTLTLOVIAAQAQREGKTCFIDAHAHALDIYARKLGVDIDNLLCSQPD	120
Db	61	I VEIYGPESSCKTTLTLOVIAAQAQREGKTCFIDAHAHALDIYARKLGVDIDNLLCSQPD	120
Qy	121	TGEQALEICDALARGAVDVIVDSVAALTPKATIEGIGSHMGLAARMMSQAMRKLKAG	180
Db	121	TGEQALEICDALARGAVDVIVDSVAALTPKATIEGIGSHMGLAARMMSQAMRKLKAG	180
Qy	181	NLKQSNLTLLIFINQIRMKIGVMFNPETTTGNGALPKFYASVRLDIRRIGAVKEGNNVYGS	240
Db	181	NLKQSNLTLLIFINQIRMKIGVMFNPETTTGNGALPKFYASVRLDIRRIGAVKEGNNVYGS	240
Qy	241	ETRVKVNKNTAAAPFKQAEFQILYGEINFGVGLVDLGKVEKLEKAGAWTSYKGEKIQG	300
Db	241	ETRVKVNKNTAAAPFKQAEFQILYGEINFGVGLVDLGKVEKLEKAGAWTSYKGEKIQG	300
Qy	301	GKANATAWLKONPETAKEIEKKVRELLSNPNSTP	335
Db	301	GKANATAWLKONPETAKEIEKKVRELLSNPNSTP	335

RESULT 2

```

US-10-733-782-5
; Sequence 5, Application US/10733782
; Publication No. US20040157248A1
; GENERAL INFORMATION:
; APPLICANT: Cox, Michael
; APPLICANT: Lusetti, Shelley
; APPLICANT: Egglar, Aimee
; TITLE OF INVENTION: RecA Mutants
; FILE REFERENCE: 960296.99501
; CURRENT APPLICATION NUMBER: US/10/733,782
; CURRENT FILING DATE: 2003-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-733-782-5

Query Match      100.0%; Score 1678; DB 4; Length 352;
Best Local Similarity 100.0%; Pred. No. 4.5e-152;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKFQKGSIMRGLGEDRSMDVETISTGSLSDIALGAGLPMGR 60
    |||||
Db 1 AIDENKOKALAAALGQIEKFQKGSIMRGLGEDRSMDVETISTGSLSDIALGAGLPMGR 60

Qy 61 IVEIYGPSSGGTTLTLQVIAAAQREGKTCAFDIAEHALDPIYARKLGVDIDLNCSPD 120
    |||||
Db 61 IVEIYGPSSGGTTLTLQVIAAAQREGKTCAFDIAEHALDPIYARKLGVDIDLNCSPD 120

Qy 121 TGEQALEICDALARSAGVDIVVDSVAALTTPKAETEIGEIGDSHMGLAARMMSQAMRKLAG 180
    |||||
Db 121 TGEQALEICDALARSAGVDIVVDSVAALTTPKAETEIGEIGDSHMGLAARMMSQAMRKLAG 180

Qy 181 NLKQSNTLLIFINQIRMKIGVMFNPETTGTGNALKFYASVRDLDIRRGAVKEGVNVS 240
    |||||
Db 181 NLKQSNTLLIFINQIRMKIGVMFNPETTGTGNALKFYASVRDLDIRRGAVKEGVNVS 240

Qy 241 ETRVKVVKNKIAPFKQAEPFIILYGGINFYELVDLGKVELIEKAGAWTSYKGEKIQ 300
    |||||
Db 241 ETRVKVVKNKIAPFKQAEPFIILYGGINFYELVDLGKVELIEKAGAWTSYKGEKIQ 300

Qy 301 GKANATAWLKNDPETAKEIEKKVRELLSNPNSTP 335
    |||||
Db 301 GKANATAWLKNDPETAKEIEKKVRELLSNPNSTP 335

RESULT 3
US-10-733-782-5
; Sequence 53319, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haelbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23

```

```
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: SEQUENCE RECOMBINATION
; CURRENT APPLICATION NUMBER: US/10/194,686
; PRIOR FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-194-686-8

Query Match          100.0%; Score 1678; DB 4; Length 358;
Best Local Similarity 100.0%; Pred. No. 4.6e-152; Indels 0; Gaps 0;
Matches 335; Conservative 0; Mismatches 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVLDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVLDNLLCSQPD 126

Qy 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIEGDSHMGSLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIEGDSHMGSLAARMMSQAMRKLKAG 186

Qy 181 NLKQSNLTLLIIFINQIRMKIGVMFGNPETTTGKNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKQSNLTLLIIFINQIRMKIGVMFGNPETTTGKNALKFYASVRLDIRRIGAVKEGENVVGS 246

Qy 241 ETRVKKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKELIEKAGAWYSYKGEKIGQ 300
Db 247 ETRVKKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKELIEKAGAWYSYKGEKIGQ 306

Qy 301 GKANATAWLKNPETAKEIEKKVRELLSNPNSTP 335
Db 307 GKANATAWLKNPETAKEIEKKVRELLSNPNSTP 341

RESULT 5
US-10-194-686-10
; Sequence 10, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: SEQUENCE RECOMBINATION
; CURRENT APPLICATION NUMBER: US/10/194,686
; PRIOR FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; OTHER INFORMATION: e. coli sequence
US-10-194-686-15
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-194-686-10

Query Match          100.0%; Score 1678; DB 4; Length 358;
Best Local Similarity 100.0%; Pred. No. 4.6e-152; Indels 0; Gaps 0;
Matches 335; Conservative 0; Mismatches 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVLDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVLDNLLCSQPD 126

Qy 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIEGDSHMGSLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIEGDSHMGSLAARMMSQAMRKLKAG 186

Qy 181 NLKQSNLTLLIIFINQIRMKIGVMFGNPETTTGKNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKQSNLTLLIIFINQIRMKIGVMFGNPETTTGKNALKFYASVRLDIRRIGAVKEGENVVGS 246

Qy 241 ETRVKKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKELIEKAGAWYSYKGEKIGQ 300
Db 247 ETRVKKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKELIEKAGAWYSYKGEKIGQ 306

Qy 301 GKANATAWLKNPETAKEIEKKVRELLSNPNSTP 335
Db 307 GKANATAWLKNPETAKEIEKKVRELLSNPNSTP 341

RESULT 6
US-10-194-686-15
; Sequence 15, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: SEQUENCE RECOMBINATION
; CURRENT APPLICATION NUMBER: US/10/194,686
; PRIOR FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; OTHER INFORMATION: e. coli sequence
US-10-194-686-15
```

Query Match 100.0%; Score 1678; DB 4; Length 358;
 Best Local Similarity 100.0%; Pred. No. 4.6e-152;
 Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
 DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 66

QY 61 IVEIYGPSSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
 DB 67 IVEIYGPSSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 126

QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
 DB 127 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 186

QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPGPPTTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
 DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPGPPTTTGGNALKFYASVRLDIRRIGAVKEGENVVG 246

QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLIIEKAGAWYSYKGEKIGQ 300
 DB 247 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLIIEKAGAWYSYKGEKIGQ 306

QY 301 GKANATAWLKNDPETAKEIEKKVRELLLSNPSTP 335
 DB 307 GKANATAWLKNDPETAKEIEKKVRELLLSNPSTP 341

RESULT 7
 US-10-733-782-3
 ; Sequence 3, Application US/10733782
 ; Publication No. US20040157248A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cox, Michael
 ; APPLICANT: Lusetti, Shelley
 ; APPLICANT: Eggett, Aimee
 ; TITLE OF INVENTION: RecA Mutants
 ; FILE REFERENCE: 960296.99501
 ; CURRENT APPLICATION NUMBER: US/10/733,782
 ; CURRENT FILING DATE: 2003-12-11
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 3
 ; LENGTH: 335
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 US-10-733-782-3

Query Match 99.8%; Score 1674; DB 4; Length 335;
 Best Local Similarity 99.7%; Pred. No. 1e-151;
 Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
 DB 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60

QY 61 IVEIYGPSSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
 DB 61 IVEIYGPSSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120

QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
 DB 121 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180

QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPGPPTTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
 DB 181 NLKQSNLTLLIFINQIRMKIGVMFGNPGPPTTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240

QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLIIEKAGAWYSYKGEKIGQ 300
 DB 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLIIEKAGAWYSYKGEKIGQ 300

QY 301 GKANATAWLKNDPETAKEIEKKVRELLLSNPSTP 335
 DB 301 GKANATAWLKNDPETAKEIEKKVRELLLSNPSTP 335

RESULT 8
 US-10-194-686-9
 ; Sequence 9, Application US/10194686
 ; Publication No. US20030148309A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DEL CARDAYRE, STEPHEN
 ; APPLICANT: TOBIN, MATTHEW
 ; APPLICANT: STEMMER, WILLEM P. C.
 ; APPLICANT: NESS, JON E.
 ; APPLICANT: MINSHULL, JEREMY
 ; APPLICANT: PATTEN, PHILLIP
 ; APPLICANT: SUBRAMANIAN, VENKITESWATAN
 ; APPLICANT: CASTLE, LINDA A.
 ; APPLICANT: KREBER, CLAU M.
 ; APPLICANT: BASS, STEVE
 ; APPLICANT: ZHANG, YING-XIN
 ; APPLICANT: COX, TONY
 ; APPLICANT: HUISMAN, GJALT
 ; APPLICANT: YUAN, LING
 ; APPLICANT: AFFHOLTER, JOSEPH A.
 ; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
 ; TITLE OF INVENTION: SEQUENCE RECOMBINATION
 ; FILE REFERENCE: 02-020730US
 ; CURRENT APPLICATION NUMBER: US/10/194,686
 ; CURRENT FILING DATE: 2002-07-11
 ; PRIOR APPLICATION NUMBER: US/09/354,922
 ; PRIOR FILING DATE: 1999-07-15
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 9
 ; LENGTH: 358
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 US-10-194-686-9

Query Match 99.7%; Score 1673; DB 4; Length 358;
 Best Local Similarity 99.4%; Pred. No. 1.4e-151;
 Matches 333; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
 DB 7 AIDENKQKALATAGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 66

QY 61 IVEIYGPSSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
 DB 67 IVEIYGPSSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 126

QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
 DB 127 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 186

QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPGPPTTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
 DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPGPPTTTGGNALKFYASVRLDIRRIGAVKEGENVVG 246

QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLIIEKAGAWYSYKGEKIGQ 300
 DB 247 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLIIEKAGAWYSYKGEKIGQ 306

QY 301 GKANATAWLKNDPETAKEIEKKVRELLLSNPSTP 335
 DB 307 GKANATAWLKNDPETAKEIEKKVRELLLSNPSTP 341

RESULT 9
 US-10-194-686-12
 ; Sequence 12, Application US/10194686


```

; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; PRIOR FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-194-686-12

Query Match          99.7%; Score 1673; DB 4; Length 358;
Best Local Similarity 99.4%; Pred. No. 1.4e-151;
Matches 333; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDMDVETITSGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDMDVETITSGSLSLDIALGAGGLPMGR 66

QY 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVLDINLLCSQPD 120
DB 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVLDINLLCSQPD 126

QY 121 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGEGISHMGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGEGISHMGLAARMMSQAMRKLKAG 186

QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246

QY 241 ETRVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVKEKLIKAGAWTSYKGEKIQ 300
DB 247 ETRVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVKEKLIKAGAWTSYKGEKIQ 306

QY 301 GKANATWLKONPETAKEIEKKVRELLSNPNSTP 335
DB 307 GKANATWLKONPETAKEIEKKVRELLSNPNSTP 341

RESULT 10
US-10-194-686-13
; Sequence 13, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN

```

```

; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; CURRENT FILING DATE: 2002-07-11
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-194-686-13

Query Match          98.9%; Score 1660; DB 4; Length 358;
Best Local Similarity 98.8%; Pred. No. 2.5e-150;
Matches 331; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDMDVETITSGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDMDVETITSGSLSLDIALGAGGLPMGR 66

QY 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVLDINLLCSQPD 120
DB 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVLDINLLCSQPD 126

QY 121 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGEGISHMGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGEGISHMGLAARMMSQAMRKLKAG 186

QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246

QY 241 ETRVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVKEKLIKAGAWTSYKGEKIQ 300
DB 247 ETRVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVKEKLIKAGAWTSYKGEKIQ 306

QY 301 GKANATWLKONPETAKEIEKKVRELLSNPNSTP 335
DB 307 GKANATWLKONPETAKEIEKKVRELLSNPNSTP 341

RESULT 11
US-10-194-686-11
; Sequence 11, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE

```

Thu Feb 16 11:15:02 2006

```
; TITLE OF INVENTION: SEQUENCE RECOMBINATION
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-194-686-11

Query Match      98.9%; Score 1659; DB 4; Length 358;
Best Local Similarity 99.1%; Pred. No. 3.1e-150;
Matches 332; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 66
QY 61 IVEIYGPSSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPSSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDVIVVDSVAALTTPKAEIEGEGIDSHMGLAARMMSQAMRKLAG 180
DB 127 TGEQALEICDALARSGAVDVIVVDSVAALTTPKAEIEGEGIDSHMGLAARMMSQAMRKLAG 186
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
DB 187 NLKLSNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 246
QY 241 ETRVKVVKNKIAAPPKQAEFQILYGEINFGYELVDLGVKEKLEIKAGAWYSYKGEKIQ 300
DB 247 ETRVKVVKNKIAAPPKQAEFQILYGEINFGYELVDLGVKEKLEIKAGAWYSYKGEKIQ 306
QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
DB 307 GKANAALWLNKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 12
US-10-282-122A-75948
; Sequence 75948, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347

; TITLE OF INVENTION: SEQUENCE RECOMBINATION
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75948
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-10-282-122A-75948

Query Match      97.8%; Score 1641; DB 4; Length 353;
Best Local Similarity 97.3%; Pred. No. 1.6e-146;
Matches 326; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
DB 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 61
QY 61 IVEIYGPSSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 62 IVEIYGPSSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 121
QY 121 TGEQALEICDALARSGAVDVIVVDSVAALTTPKAEIEGEGIDSHMGLAARMMSQAMRKLAG 180
DB 122 TGEQALEICDALARSGAVDVIVVDSVAALTTPKAEIEGEGIDSHMGLAARMMSQAMRKLAG 181
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
DB 182 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 241
QY 241 ETRVKVVKNKIAAPPKQAEFQILYGEINFGYELVDLGVKEKLEIKAGAWYSYKGEKIQ 300
DB 242 ETRVKVVKNKIAAPPKQAEFQILYGEINFGYELVDLGVKEKLEIKAGAWYSYKGEKIQ 301
QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
DB 302 GKANATTLKONPETAKEIEKKVRELLLSNPSTP 336

RESULT 13
US-10-282-122A-59525
; Sequence 59525, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
```

```
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59525
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-10-282-122A-59525

Query Match          96.1%; Score 1613; DB 4; Length 352;
Best Local Similarity 95.5%; Pred. No. 7, 7e-146;
Matches 320; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALQIEKQFGKGSIMRLGEDRSDMDVETISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALQIEKQFGKGSIMRLGEDRSDMDVETISTGSLSLDIALGAGGLPMGR 61
QY 61 IVEIYGPSSSGKTTLTQVIAAAQREGKTCAFIDAEHALDPYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSSGKTTLTQVIAAAQREGKTCAFIDAEHALDPYARKLGVDIDNLLCSQPD 121
QY 121 TGEQALEICDALARSGVDVIVDSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALARSGVDVIVDSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 181
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGNNVVG 240
Db 182 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGNNVVG 241
QY 241 ETRVKVKNKTAAPFKQAEFOILYCEGINFYGELVDLGVKELIEKAGAWTSYKGEKIQG 300
Db 242 ETRVKVKNKTAAPFKQAEFOILYCEGINFYGELVDLGVKELIEKAGAWTSYKGEKIQG 301
QY 301 GKANATWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 302 GKANAITWLENPAARAKEIEKKVRELLLNQDQAKP 336

RESULT 14
US-10-282-122A-55632
; Sequence 55632, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
```

```
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55632
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-10-282-122A-55632

Query Match          94.9%; Score 1593; DB 4; Length 352;
Best Local Similarity 94.6%; Pred. No. 6, 4e-144;
Matches 317; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALQIEKQFGKGSIMRLGEDRSDMDVETISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALQIEKQFGKGSIMRLGEDRSDMDVETISTGSLSLDIALGAGGLPMGR 61
QY 61 IVEIYGPSSSGKTTLTQVIAAAQREGKTCAFIDAEHALDPYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSSGKTTLTQVIAAAQREGKTCAFIDAEHALDPYARKLGVDIDNLLCSQPD 121
QY 121 TGEQALEICDALARSGVDVIVDSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALARSGVDVIVDSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 181
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGNNVVG 240
Db 182 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGNNVVG 241
QY 241 ETRVKVKNKTAAPFKQAEFOILYCEGINFYGELVDLGVKELIEKAGAWTSYKGEKIQG 300
Db 242 ETRVKVKNKTAAPFKQAEFOILYCEGINFYGELVDLGVKELIEKAGAWTSYKGEKIQG 301
QY 301 GKANATWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 302 GKANAISWLENPAARAKEIEKKVRELLLNQDQSKP 336

RESULT 15
US-10-282-122A-78524
; Sequence 78524, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
```

Thu Feb 16 11:15:02 2006

```

; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 78524
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Yersinia pestis
US-10-282-122A-78524

Query Match          91.4%; Score 1534; DB 4; Length 356;
Best Local Similarity 91.9%; Pred. No. 3e-138;
Matches 305; Conservative 14; Mismatches 13; Indels 0; Gaps 0;

QY      1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDEVETISTGSLSDIALGAGGLPMGR 60
Db      2 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDEVETISTGSLSDIALGAGGLPMGR 61

QY      61 IVEIYGPSSGKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db      62 IVEIYGPSSGKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 121

QY      121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKL 180
Db      122 TGEQALEICDALTFSGNVDVIIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKL 181

QY      181 NLKQSNLTLLIFINQIRMKIGWFGNPGTTTGGNALKFYASVRLDIRRIGAVKEGENVWGS 240
Db      182 NLKANTLLIFINQIRMKIGWFGNPGTTTGGNALKFYASVRLDIRRIGAVKDGVDVWGS 241

QY      241 ETRVKVYKVKIAAPFKQAEQIILYGEINFGYELVDLGVEKLEIEKAGAWYSYKGEKIGQ 300
Db      242 ETRVKVYKVKIAAPFKQAEQIILYGEINFGYELVDLGVEKLEIEKAGAWYSYKGEKIGQ 301

QY      301 GKANATAWKNDPETAKEIEKKVRELLLSNPN 332
Db      302 GKANASNYLKENPAIAAELOKUREMLLNGN 333
```

Search completed: February 16, 2006, 01:44:34
Job time : 124 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 01:40:35 ; Search time 10.5 Seconds
(without alignments)
453.421 Million cell updates/sec

Title: US-10-733-782-1
Perfect score: 1678
Sequence: 1 AIDENKQKALAAALQIEKQ.....AKEIEKKVRELLLSNPNTSP 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 107799 seqs, 14211699 residues

Total number of hits satisfying chosen parameters: 107799

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_New.*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1212.5	72.3	375	6	US-10-467-657-4766	Sequence 4766, Ap
2	1019	60.7	363	7	US-11-074-176-296	Sequence 296, App
3	970.5	57.8	355	7	US-11-098-686-11427	Sequence 11427, A
4	155	9.2	443	7	US-11-098-686-10861	Sequence 10861, A
5	111.5	6.6	459	6	US-10-467-657-3092	Sequence 3092, A
6	109	6.5	453	6	US-10-467-657-206	Sequence 206, App
7	109	6.5	453	6	US-10-467-657-6400	Sequence 6400, Ap
8	104	6.2	454	7	US-11-098-686-10599	Sequence 10599, A
9	102	6.1	453	6	US-10-467-657-3626	Sequence 3626, Ap
10	101	6.0	453	6	US-10-454-437-50	Sequence 50, Appl
11	100	6.0	471	6	US-10-467-657-802	Sequence 802, App
12	99	5.9	421	6	US-10-467-657-4148	Sequence 4148, Ap
13	97	5.8	365	7	US-11-082-389-132	Sequence 132, App
14	97	5.8	376	7	US-11-213-368-14	Sequence 14, App
15	97	5.8	479	7	US-11-082-389-130	Sequence 130, App
16	96	5.7	286	7	US-11-082-389-88	Sequence 88, Appl
17	95.5	5.7	419	6	US-10-979-821-6	Sequence 6, Appli
18	95.5	5.7	419	7	US-11-114-922-6	Sequence 6, Appli
19	95	5.7	257	7	US-11-156-084-264	Sequence 264, App
20	95	5.7	376	7	US-11-082-389-122	Sequence 122, App
21	95	5.7	376	7	US-11-082-389-124	Sequence 124, App
22	94.5	5.6	640	7	US-11-205-109-9	Sequence 9, Appli
23	94.5	5.6	716	6	US-10-131-826A-96	Sequence 96, Appl
24	94	5.6	406	6	US-10-467-657-7420	Sequence 7420, Ap
25	94	5.6	655	7	US-11-094-586-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-10-467-657-4766
; Sequence 4766, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 4766
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
; US-10-467-657-4766

Query Match	72.3%	Score 1212.5;	DB 6;	Length 375;
Best Local Similarity	69.0%	Pred. No. 4e-86;		
Matches	225;	Conservative 52;	Mismatches 48;	Indels 1; Gaps 1;
Qy	2	IDENKQKALAAALQIEKQFGKGSIMRL-GEGRSDMDVETISTGSLSLDIALGAGGLPMGR	60	
Db	28	MSDDSKALAAALQIEKSGKGAIMKMDGSGQEENLEVIETGSLGLDIALGVGGLPRGR	87	
Qy	61	IVEIYGPSSGKTTLTQVIAAQAQREGKTCAFIDAEHALDPIYARKLGVDNLLCSQPD	120	
Db	88	IVEIYGPSSGKTTLCLEAVAAQCKNGGVCAFPVDAEHAFDPVYARKLGKVKVEELYLSQPD	147	
Qy	121	TGEQALEICDALARAGVDVIVDVSVAALTPKAIEGEGIGSHMGLAARMMSQAMRKLKAG	180	
Db	148	TGEQALEICDVLVSAGGIDVVDVSVAALTVPKAEIEMDGMDSHVGQLARLMSQALRKLKG	207	
Qy	181	NLKQSTLLIPINQIRMKIGVMFGNPETTTGNAKLFYASVRLDIRRTGAVKEGNNVGS	240	
Db	208	HIKKTNTLVVFQINQIRMKIGVMFGSPETTTGNAKLFYSSVRLDIRRTGSIKKGSEVLGN	267	
Qy	241	ETRVKVNKNTAAPKQAEFOILYEGEINFYGVLDLGVKELTEKAGAWSYKGEKIGQ	300	
Db	268	ETRVKVNKNVAPPFRQAEFDILYEGEISWEGELIDIGVKNDIINKSGAWSYNGAKITG	327	

Thu Feb 16 11:15:02 2006

```
QY 301 GKANATAWLKONPETAKEIEKKVREL 326
DB 328 GKDNVRVWLKENPEISDEIDAKIRAL 353

RESULT 2
US-11-074-176-296
; Sequence 296, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: McAuliffe, Eric
; APPLICANT: Peril, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 296
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-296

Query Match 60.7%; Score 1019; DB 7; Length 363;
Best Local Similarity 56.8%; Pred. No. 2.8e-71;
Matches 183; Conservative 74; Mismatches 65; Indels 0; Gaps 0;

QY 4 ENKQKALAAALQIEKQFGKSGIMRLGEDRSMVDVETISGSLDIALGAGGLPMGRIVE 63
DB 4 DEKKAALDALKKIEKNFGKSGAVMRMGKADTQISTVPTGSLDAAIGVGGYPRGRIIE 63

QY 64 IYGPSSGKTTLTQVIAAAREGKTCFAIDAEHALDPIYARKLVGDIDNLCSQPDGTG 123
DB 64 IYGPSSGKTTLTQVIAAAREGKTCFAIDAEHALDPIYARKLVGDIDNLCSQPDGTG 123

QY 124 QALEICDALARSGADVIVVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLAGNLK 183
DB 124 EGIQIADTLSSGAIDIVVDSVAALVPRAEIEGEGMDAHVGLQARLMSQALRKLSTGIS 183

QY 184 QSNLLIFINQIRMKIGVMFGNPTETTTGGNALKFYASVRLDIRRIGAVKEGENVVGSETR 243
DB 184 KTKTIAIFINQIREKVGVMFGNPTETTPGRALKFYSTVRLEVRRAEQIKSGDVLGNRVK 243

QY 244 VKVVKNKIAAPFKQAEFOILYGEINFGYBELVDLVGKKEKLIKAGAWYSYKGEKIGOGKA 303
DB 244 IKVVKNKVAPFKVAEVDIMYKGISGOSGELLNAAKDIDKAGSWYSYKSDRIGOGRE 303

QY 304 NATAWLKNDPETAKEIEKKVRE 325
DB 304 NAKKYLEEHPDIYQKVQEQVRQ 325

RESULT 3
US-11-098-686-11427
; Sequence 11427, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
```

```
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11427
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-11427

Query Match 57.8%; Score 970.5; DB 7; Length 355;
Best Local Similarity 57.3%; Pred. No. 1.5e-67;
Matches 189; Conservative 64; Mismatches 76; Indels 1; Gaps 1;

QY 3 DENKQKALAAALQIEKQFGKSGIMRLGEDRSMVDVETISGSLDIALGAGGLPMGRIV 62
DB 9 EDSREALKTALDTIERKFGGAVMKLSDDVHVHVAVIPTGSGIGLJDLALGIGGIPRGRTV 68

QY 63 EYGPSSGKTTLTQVIAAAREGKTCFAIDAEHALDPIYARKLVGDIDNLCSQPDGTG 122
DB 69 EYGPSSGKTTLTQVIAAAREGKTCFAIDAEHALDPIYARKLVGDIDNLCSQPDGTG 128

QY 123 EQALEICDALARSGADVIVVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLAGNL 182
DB 129 EQALEIADMLVRSAGVDLVVDSVAALTPQTELESGMGETQVGGHARLMSHALRKLGTGT 188

QY 183 KQSNLLIFINQIRMKIGVM-FGNPTETTTGGNALKFYASVRLDIRRIGAVKEGENVVGSE 241
DB 189 HKSHAVTAVIFINQIRMKIGVM-FGNPTETTTGGNALKFYASVRLDIRRIGAVKEGENVVGSE 248

QY 242 TRKVVKNKIAAPFKQAEFOILYGEINFGYBELVDLVGKKEKLIKAGAWYSYKGEKIGOG 301
DB 249 TRKVVKNKIAAPFKQAEFOILYGEINFGYBELVDLVGKKEKLIKAGAWYSYKGEKIGOG 308

QY 302 KANATAWLKNDPETAKEIEKKVRELLSNP 331
DB 309 KEKVRALLDENTSLKSSITETALMEHLGNP 338

RESULT 4
US-11-098-686-10861
; Sequence 10861, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10861
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10861

Query Match 9.2%; Score 155; DB 7; Length 443;
Best Local Similarity 32.1%; Pred. No. 0.0001;
Matches 62; Conservative 28; Mismatches 81; Indels 22; Gaps 8;
```

```
QY 18 EKQFGKSGIMRLGEDRSMVDVETISGSLDIALGAGGLPMGRIVEIYGPSSGKTTLT 77
DB 44 QKIYKKLKPPIGRVETSTYQPFSTGIVKLDHLGKLVFSSALL-IGGEPGIGKSTLL 102

QY 78 QVIAAAREGKTCFAIDAEHALDPIYAR--KLGVVDNLCSQPDGTGEQALEICDALARS 135
DB 103 QLAGATATLCKIVLYVSGEESLPQIKARABRLVTLHDNLISM---ATSQVEDILPLNTS 159
```

QY 136 GAVDVIVDVSAALTPKABIEGIGDSHMGSLAARMMSQAMRKLAGNL-----KQSNLTLLIP 191
DB 160 EAPDLLLDSVQTLT-SSNAELPGN-----VSQ-VRAVATELVBEACQKGTTVIF 208
QY 192 INQIRMKIGVMFG 204
DB 209 IGHV-TKDGTLAG 220

RESULT 5

US-10-467-657-3092
; Sequence 3092, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 3092
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3092

Query Match 6.6%; Score 111.5; DB 6; Length 459;

Best Local Similarity 22.0%; Pred. No. 0.23;
Matches 65; Conservative 44; Mismatches 127; Indels 59; Gaps 12;
QY 25 SIMRIGEDRSMVDVETISTGSLSLDIALGAGGLPMGRIVEIYGPSSGKTTTLQVIAAAQ 84
DB 57 TVQSLSAVTAAEVPRNPTGMGELDRVLG-GGLVDGAVILLGGDPGIGKSTLLIQTIAKWA 115
QY 85 REGKTCATDAEHALDPYARKLVGVDID-----NLCSQPDGTGEQALEICDALARSGVDV 140
DB 116 QSRKV-LVVSSESAQQVALRSQRLEHAEGVNNLL-----ABIRMEAIQRAALKQHQPEV 168
QY 141 IWDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLAGNL-----KQSNLTLLIPNOIR 196
DB 169 VVIDSIQWY-----SDQITSAPGSVSQ-VRECAQLTRWAKQWGMILVGHV- 216
QY 197 MKIGVMGNPETTT-----GGNALKFYASVRLDIRRIGAVKEGENVVVGSETRVKVW 247
DB 217 TKDGAIAQ-PRVLEHMDVTLVYFEGDQHSNYRMIRAIKNRFGAANELGVFAMTENGAKGV 275
QY 248 KNTIA---APFKQ-----ABFOILYGEGINFYGELVDLGKVE 281
DB 276 SNPSAIFLASYRDDTPFGSCVLVTQBSRPLLVIEIQALVDDAHGFTPKELSVGLEQ 330

RESULT 6

US-10-467-657-206
; Sequence 206, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8

; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 206
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-206

Query Match 6.5%; Score 109; DB 6; Length 453;

Best Local Similarity 23.5%; Pred. No. 0.36;
Matches 65; Conservative 43; Mismatches 91; Indels 78; Gaps 16;
QY 1 AIDENKOKALAAALGQ-IEKQFGKGSIMRLGEDRSMVDVET--ISTGSLSLDIALG-AGGL 56
DB 151 AVGKDAVKRETKTQGTVEDLIG-----GLDKELDGVRFGLPTGLMKLD---GWTGGL 200
QY 57 PMGRIVEIYGPSSGKTTTLTQVIAAAREGKTCAPIDAHALDPI----YARKLGVDDID 112
DB 201 PDGNLIVIAARPSMGKTVLAENIARFALKQKQKAVHFQSYEMSAVELARRGMAECNIPMQ 260
QY 113 NL-----LCSPDGTGEQALE-----ICDAL-----ARSGAVDVIVVDVSV 146
DB 261 NLKTNLTQSDYANNPIYVSQAKWKFDVNCDDLNVDELCLFLAKEKKLTTGLDILLVDHL 320
QY 147 AALTPKA--EIEGEIGDSHMGSLAARMMSQAMRKLAGNLKQSNLTLLIFINQIRMKIGVMFG 204
DB 321 -HIMPRAGRDEVAELGN-----ISRRLKNLAEL---NTPVVLVAQLNR-----G 361
QY 205 NPETTTGNALKFYASVR---LDIRRIGAVKEGENVV 238
DB 362 NTK-----QADKEPNMADIRGSAIEQDANII 388

RESULT 7

US-10-467-657-6400
; Sequence 6400, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 6400
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6400

Query Match 6.5%; Score 109; DB 6; Length 453;

Best Local Similarity 23.5%; Pred. No. 0.36;
Matches 65; Conservative 43; Mismatches 91; Indels 78; Gaps 16;
QY 1 AIDENKOKALAAALGQ-IEKQFGKGSIMRLGEDRSMVDVET--ISTGSLSLDIALG-AGGL 56
DB 151 AVGKDAVKRETKTQGTVEDLIG-----GLDKELDGVRFGLPTGLMKLD---GWTGGL 200
QY 57 PMGRIVEIYGPSSGKTTTLTQVIAAAREGKTCAPIDAHALDPI----YARKLGVDDID 112
DB 201 PDGNLIVIAARPSMGKTVLAENIARFALKQKQKAVHFQSYEMSAVELARRGMAECNIPMQ 260
QY 113 NL-----LCSPDGTGEQALE-----ICDAL-----ARSGAVDVIVVDVSV 146
DB 261 NLKTNLTQSDYANNPIYVSQAKWKFDVNCDDLNVDELCLFLAKEKKLTTGLDILLVDHL 320

Qy	147	AALTFKA--ETEGEIGD	SHGLAARMMSQ	AMRKLAGNLKQ	SNTLLIFINQIR	MTKIVMFG	204
		: : :	: : :	: : :	: : :	: : :	
Db	321	HINPRGRD	VEAELGN-----	ISRLKLAEL--	NTPVVLAQ	LNR-----	361
Qy	205	NPETTTG	NALKFVASVR--	LDIRRICAV	KEGVNV		238
		: : :	: : :	: : :	: : :		
Db	362	NTK-----	QADKEPN	MADIRGSA	IEODANII		388

```

RESULT 8
US-11-098-686-10599
; Sequence 10599, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10599
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10599

```

Query Match	6.2%	Score 104;	DB 7;	Length 454;	
Best Local Similarity	22.0%;	Pred. No. 0.86;			
Matches	55;	Conservative 45;	Mismatches 94;	Indels 56;	Gaps 11;
Qy	16	QIEQKFGKSIWRLGEDRSMDEVETISGSLSDIALNGGLPMGRIVEIYGPESGKTTL	75		
Db	168	ELEKRF-----ERKEQVTGTTGYNRLD-KLTAGLQPSDLIIIVAARP-SWGKTAF	215		
Qy	76	TLOW-IAAARGKTKAFIDAEHALDPYARKL-----GVDDINLLCS--QPDTEQOALE	128		
Db	216	SLNMAWRSAIQOQTPVALYISLMSMNQLMMRMLCANGKVDLSHLAHGYLNSDEWSRLYHA	275		
Qy	129	CDALARGSAVDVIVVDSVAALTP-----KAETEGE-----GDSHMGAA	168		
Db	276	ADVLGQA----PIFIDDPALSPLELRARTRLKIESDVGLVWDYDLQMRGNKRTDSRE	331		
Qy	169	RMSQAMRKLAGNLKOSNTLIFINQIRMKIGWMPGNPETTTGGNALKFYASVRLDIRRI	228		
Db	332	QEISEISRLSLAKEINIPVVALSQLNKL-----EDRTDKRPQLS-----DLRES	378		
Qy	229	GAVKEGENVV	238		
Db	379	GAIEQDAQVI	388		

RESULT 9
US-10-467-657-3G26
; Sequence 3626, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8

```

; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 3626
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3626

      Query Match      6.1%; Score 102; DB 6; Length 453;
      Best Local Similarity 23.4%; Pred. No. 1.2;
      Matches 68; Conservative 42; Mismatches 87; Indels 94; Gaps

Qy      2 IDENKOKA---LAAALGQIEKQFGKGSIMRL-----GEDRSMDVET--IST
Db      138 VAEKLSKAADELAAA-----GKDAVKRETKTFTGQTVLEDIGGLDKRLDGVRFGLPT
Qy      44 SLSLDITAGL-AGGLPMGRIVEIYGPSSSGKTTTLTQVIAAQRGKTCFIDAEHALDPP
Db      190 LMKLD---GWTGGLPDGNLIVIAARPSMGKTVLAENIARFALKQKAVHFQSYEMSAVE
Qy      103 ---YARKLGVDIDNL-----LCSQPDTEQQALE-----ICDAL
Db      247 ARRGMAAECNI PMONLKTGNLTOSDYANMFIYVSQAKEMKFDVNCIDLINVDLCLFLAKE
Qy      133 ARGAVDVIYVDSVAALTPKA--EIEGEIGDSHMGLAARMMSQAMRKLAGNLKQSNLTLL
Db      307 KLTGGLDLLLVVDHL-HMFPAGRDVEALGN-----ISRLKNLAEL---NIPV
Qy      191 FNIQRMKIGVMFGNPTTTGGNALPKFYASVR---LDIRIGAVKEGVNV 238
Db      354 LVLAOLNR-----GNTK-----QADKRPNMADIRGSGAIEODANII 368

```

RESULT 10
US-10-454-437-50
; Sequence 50, Application US/10454437
; Publication No. US2005027115A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION
; FILE REFERENCE: BGI-128CFCN
; CURRENT APPLICATION NUMBER: US/10/454,437
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 50
; LENGTH: 925


```
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-454-437-50

Query Match
Best Local Similarity 6.0%; Score 101; DB 6; Length 925;
Matches 75; Conservative 67; Mismatches 125; Indels 120; Gaps 18;

Qy 14 LQOIBKQFGKGS--IMRLGED-----RSMDEVETIS-----TGLSLDIALGAG 54
Db 110 LGLIREGEGVAAQVLKLGADLPVRQVQVQLLSGYEGGQSPGQGGQAPGCGAVGAG 169
Qy 55 GLPMGRIVEIYGPSSG-----KTTTLT-----QVIAAAREKTCFADAEHALDPIYAR 105
Db 170 AAPGGR-----PSSGSPGERSTSLVLDQFGRNLTQAAKGK-----LDPVVR 212
Qy 106 KLGVD-----IDNLLCSQDPTGEOALEICDALARSADVIVVDSVAALTPKAE 154
Db 213 DKEIRIMQVLSRRTPKNNPVLIGEPGVGTAV-----VEGLADIVNGKVPETL 261
Qy 155 IEGEIGDSHMG-LAARM-----MSQAMRKLAGNLKQSNLTLLIFINQIRMKIGVMFGNPET 208
Db 262 KDKQVYSLDGLSLGAVSRVGRDFEERLKKVLKEINQRGDIILFIDEIHTLVGA--GAERG 319
Qy 209 TTGNALKFYASVRLDIRIGAV-----KEGENVVGSETRVKKVKNKIAAPFKQAEFQI 262
Db 320 AIDAAASLLKPLARGELQTIATTLDEYRKHKIEKDAALERRFPV--QVPEPSVDLTVEI 377
Qy 263 LYG-----EGINFYGELVDLGVEK-----LIEKAGAWSYKGEKIG 299
Db 378 LKGLRDRVEAHRVSIITGALTAAQLADRYINDRFLPDKAVDLIDEAGARMRIK----- 432
Qy 300 QCKANATAWLKONPETAKEIEKKVREL 326
Db 433 -----RMTA-----PSSLREVDRIADV 450

RESULT 11
US-10-467-657-802
; Sequence 802, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 802
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-802

Query Match 6.0%; Score 100; DB 6; Length 471;
Best Local Similarity 25.8%; Pred. No. 1.8;
Matches 80; Conservative 40; Mismatches 128; Indels 62; Gaps 17;

Qy 25 SIMRLGEDRSMDEVETIST-----GSLSLDIALGAGGLPMGRIVEIYGPSSGKTTLT 76
Db 113 TLVLYGVPLIDVETLETLLLEAAGNEVGLLTDVPADPAGL--GRIIR-----DGSQSVTAI 166
Qy 77 LQVIAAAREKTCFADAEHALDPIYARKLVGDINDNLLCSQDPTGEOALEICDALARS 136
Db 167 VEEKDASATQ-KTIREINTGLVLP-----NAKLENWLSLSSNNAQGEYLLDTLIAKA 219

; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-454-437-50

Query Match
Best Local Similarity 6.0%; Score 99; DB 6; Length 291;
Matches 63; Conservative 44; Mismatches 83; Indels 76; Gaps 15;

Qy 55 GLPM--GRIVEIYGPSSGKTTLTLOVIAAAREKTCFAD--AEHALDPIYARKLV 109
Db 70 GIPLMNRQKVIAIDPGASGKGTVAARVAAALGYD-----YLDTCALYRLTALYAKQGV 124
Qy 110 --DIDNLLCSQDPTGEOALEICDALAR-----SGAVDVIVVDSVAALTPKAEIGEIG 160
Db 125 EWHDEENV-----SALAKLPAVESG--NRILLD-----EDVSDGIR 160
Qy 161 DSHMGLAARMMSQAMRKLAGNLKOSNTLLI---FINQIRMKIGVMFGNPETTTGNALKF 217
Db 161 TEAIGMGASAVAQWPKVRAALLQQRDFLTKGLVADGRDTGVSVVFPQAE-----LKI 213
Qy 218 YASVRLDIR-----RIGAVKEG---ENVVGS-ETRVKVKVKNKIAAPFKQAEFQILYCEG 267
Db 214 FLTAESKIRARRAKQIGICEGTFERILSDIETRDADRNRKRVAPLKQOPDAL----- 268
Qy 268 INFYGELVDLG--VKEKLEIKAGAWY 291
Db 269 -----LLDTSRLTIETVKKVLDWY 288

RESULT 13
US-11-082-389-132
; Sequence 132, Application US/11082389
; Publication No. US20050244935A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
```

```

; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberman, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-131PCPCN
; CURRENT APPLICATION NUMBER: US/11/082,389
; PRIOR FILING DATE: 2005-03-16
; PRIOR APPLICATION NUMBER: US 09/603024
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/143262
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: US 60/151281
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930487.4
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19930489.0
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931549.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931550.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932134.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 446
; SEQ ID NO 132
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-082-389-132

Query Match 5.8%; Score 97; DB 7; Length 365;
Best Local Similarity 20.5%; Pred. No. 2.2;
Matches 67; Conservative 54; Mismatches 126; Indels 80; Gaps 12;

QY 7 QKALAAAL---GQIEKQFGKGSIMRLGEDRSMDVETISTGSLDIALGAGGLPMGRIVE 63
Db 114 QQLDASILDQPEIASDGSAPVIDVEASKSFETALHKVSLAVR-----KGLDLG 166

QY 64 IYGPSSGKTTLTQVIAAAQREGKTCFAIDAEHALDPIYARKLGVDIDNLLCSQDPTGE 123
Db 167 IVGGSGSGKTTL-LKLI-----GLD-----KPTTG- 191

QY 124 QALEICDALARSGAVDVIVVDSVAALTPKAEIEGIDSHMGLAARMMSQAMKRLAGNLK 183
Db 192 -----TVAVTGVQVWFQDPOSSLNPRMKIDIVAEPLLGWNA-----AE 231

QY 184 QSNLTLLIFINQIRMKIGMFGNPTTTGGNALKFYASVRDIDIRIGAVKEGENVVGSETR 243
Db 232 KTRFRAEVIQVGLSPDLRYHPFSGGQ-----RQISIALALAKPA--ILLADEP 283

QY 244 KVVKVKIAFFQKAEQILYEGINIFYGELVDLVKVEKLEKAGAWSYKGEKIGOGKA 303
Db 284 VSALDVSVRKQVLDLQQLVBEYGITLVFVSHDLAVVRHLIC--TTVWVMEQGRVLEGGPI 341

QY 304 NATAWLKDNPET--AKEIEKKVRELL 328
Db 342 DS----VYDHPQETTKELLDVAPRLSL 365

```

RESULT 14

```

US-11-213-368-14
; Sequence 14, Application US/11213368
; Publication No. US20060003936A1
; GENERAL INFORMATION:
; APPLICANT: MayerKuckuk, Phillip
; APPLICANT: Banerjee, Debabrata

```

```

; APPLICANT: Bertino, Joseph R.
; TITLE OF INVENTION: Method for Modulating the Production of a Selected
; TITLE OF INVENTION: Protein In
; TITLE OF INVENTION: Vivo
; FILE REFERENCE: MSK.P-053
; CURRENT APPLICATION NUMBER: US/11/213,368
; CURRENT FILING DATE: 2005-08-26
; PRIOR APPLICATION NUMBER: US/10/421,285
; PRIOR FILING DATE: 2003-04-22
; PRIOR APPLICATION NUMBER: US 60/375,250
; PRIOR FILING DATE: 2002-04-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 376
; TYPE: PRT
; ORGANISM: herpes virus
US-11-213-368-14

Query Match 5.8%; Score 97; DB 7; Length 376;
Best Local Similarity 23.6%; Pred. No. 2.3;
Matches 53; Conservative 30; Mismatches 80; Indels 62; Gaps 11;

QY 58 MGRIVEIY--GPSSGKTTLTQVIAAAQREGKTCFAIDAEHALDPI-YARKLGVD--ID 112
Db 46 MPTLLRVYIDGPHGMGKTTTTQLLVALGSRD-----DIVVVPDPMYQVVLGASSETIA 98

QY 113 NLLCSQP-----DTGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGIDSHM- 164
Db 99 NIYTTQHRLDQGEISAGDAAVMTSAQITMG-MPYAVTDAVLA-----PHIGGEAGSSHAP 153

QY 165 -----GLAARMMSQAMKRLAGNLKQSNLTLLIFINQIRMKIGMFGNPTTTGGN 213
Db 154 PPALTLLIFDRHPDIAALLCYPAARYLMGSM-TPQAVLAFVALL-----PPTLPGTN 202

QY 214 -----ALKFYASVRDIDIRIGAVKEGENVVGSETR 243
Db 203 IVLGALPEDRHIDRLAKRQRPGERLDLMLAAIRRVYGLLANTVR 247

RESULT 15
US-11-082-389-130
; Sequence 130, Application US/11082389
; Publication No. US20050244935A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberman, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-131PCPCN
; CURRENT APPLICATION NUMBER: US/11/082,389
; CURRENT FILING DATE: 2005-03-16
; PRIOR APPLICATION NUMBER: US 09/603024
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/143262
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: US 60/151281
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930487.4
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19930489.0
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931549.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931550.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932134.5

```

```
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 446
; SEQ ID NO 130
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-11-082-389-130

Query Match      5.8%; Score 97; DB 7; Length 479;
Best Local Similarity 20.5%; Pred. No. 3.2;
Matches 67; Conservative 54; Mismatches 126; Indels 80; Gaps 12;

QY 7 QKALAAAL---GQIEKQFGKGSIMRLGEDRSMDVETISTGSLSDIALGAGGLPMGRIVE 63
Db 228 QQLLDASILDQPEIASDSGAPVVIDVEEASKSFETTALHKVSLAVR-----KGDLLG 280

QY 64 IYGPESGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPDGTGE 123
Db 281 IVGGSGGKTTL-LKLIA-----GLD-----KPTTG- 305

QY 124 QALEICDALARGAVDIVVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLGNLK 183
Db 306 -----TVAVTGGVQMVFPQDPOSSLNPRMKIKDIVAEPLLGWNA-----AE 345

QY 184 QSNTHLIFINQIRMKIGVMFGNPETTGGNALKFYASVRLDIRRIGAVKEGENVVGSETR 243
Db 346 KTRVAEVITQVGLSPDLDRYPHEFSGQ-----RQISIAIALAIKPA--ILLADEP 397

QY 244 VKVVKNKIAAPFKQAEFQILYGEINFYGELVDLGVKKEKLIKAGAWSYKGEKIQGQKA 303
Db 398 VSALDVSVRKQVLDLQQLVEEYGITLVFVSHDLAVVRHLC--TTVWVMEQGRVLEQGP 455

QY 304 NATAMLKONPET--AKEIEKKVRELLJ 328
Db 456 DS---VYDHPQTEYTKELLDVAPRLSL 479

Search completed: February 16, 2006, 01:45:01
Job time : 11.5 secs
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 01:14:10 ; Search time 132 Seconds
(without alignments)
1115.090 Million cell updates/sec

Title: US-10-733-782-3
Perfect score: 1678
Sequence: 1 AIDENKQKALAAALGQIEKQ.....AKBIEKKVRELLSNPNSTP 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*
9: Geneseqp2005s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1678	100.0	335	8	Adrl6430 E. coli R
2	1674	99.8	335	8	Adrl6428 E. coli R
3	1674	99.8	352	8	Adrl6432 E. coli w
4	1674	99.8	353	6	Abui5395 Protein e
5	1674	99.8	358	3	AAy68827
6	1674	99.8	358	3	AAy68829
7	1674	99.8	358	8	Adq87824 E. coli r
8	1674	99.8	358	8	Adq87822 E. coli r
9	1674	99.8	358	8	Adq87821 E. coli r
10	1673	99.7	358	3	AAy68831
11	1670	99.5	358	2	AAw64213
12	1669	99.5	358	3	AAy68828
13	1669	99.5	358	8	Adq87823
14	1665	99.2	358	8	Adq87826
15	1664	99.2	358	2	AAw64215
16	1661	99.0	358	2	AAw64214
17	1658	98.8	358	2	AAw64217
18	1658	98.8	358	2	AAw64218
19	1656	98.7	358	3	AAy68832
20	1656	98.7	358	8	Adq87827
21	1655	98.6	358	3	AAy68830
22	1655	98.6	358	8	Adq87825
23	1639	97.7	358	2	AAw64216
24	1637	97.6	353	6	Abu48024 Protein e

25	1609	95.9	352	6	ABU31601	Abu31601 Protein e
26	1609	95.9	356	7	ABO63433	ABO63433 Klebsiell
27	1589	94.7	352	6	ABU27708	Abu27708 Protein e
28	1543	92.0	356	6	ABM67793	Abm67793 Phototrab
29	1530	91.2	356	6	ABU50600	Abu50600 Protein e
30	1500	89.4	355	6	ABU40669	Abu40669 Protein e
31	1500	89.4	367	7	ADF05150	Adf05150 Bacterial
32	1463	87.2	351	6	ABU44971	Abu44971 Protein e
33	1432	85.3	412	6	ABU49151	Abu49151 Protein e
34	1349	80.4	354	6	ABU39491	Abu39491 Protein e
35	1328	79.1	354	6	ABU30311	Abu30311 Protein e
36	1269	75.6	398	7	ABO72381	ABO72381 Pseudomon
37	1268	75.6	346	6	ABU38581	Abu38581 Protein e
38	1257	74.9	355	6	ABU40018	Abu40018 Protein e
39	1255	74.8	349	6	ABU16844	Abu16844 Protein e
40	1255	74.8	355	6	ADA36276	Ada36276 Acinetoba
41	1255	74.8	369	6	ABU41878	Abu41878 Protein e
42	1247.5	74.3	366	8	ADL05025	Adl05025 M. catarr
43	1242.5	74.0	349	6	ABU35468	Abu35468 Protein e
44	1238.5	73.8	342	8	ADR21929	Adr21929 Methyloph
45	1237.5	73.7	348	6	ABU33125	Abu33125 Protein e

ALIGNMENTS

RESULT 1
ADRL6430
ID ADRL6430 standard; protein; 335 AA.
XX
AC ADRL6430;
XX
DT 04-NOV-2004 (first entry)
XX
DE E. coli RecA mutant protein #2.
XX
KW RecA; DNA binding protein; ssDNA binding protein; SSB;
KW DNA strand exchange reaction; site-specific cleavage; mutant; mutuin.
XX
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT Misc-difference 38
FT /note= "Wild-type Glu substituted by Lys"
XX
XX US2004157248-A1.
XX
PD 12-AUG-2004.
XX
XX 11-DEC-2003; 2003US-00733782.
XX
PR 12-DEC-2002; 2002US-0432758P.
XX
XX (COXM)/ COX M M.
XX (LUSE)/ LUSSETTI S L.
XX (EGGL)/ EGGLER A L.
XX (HARU)/ HARUTA N.
XX
XX Cox MM, Lusetti SL, Eggler AL, Haruta N;
XX
XX WPI; 2004-580265/56.
XX N-PSDB; ADRL6431.
XX
XX New RecA mutant proteins comprising a single mutation or a double
XX mutation, useful for catalyzing homologous DNA pairing and DNA strand
XX exchange reactions in an in vitro or in vivo environment.
XX
XX Claim 13; SEQ ID NO 3; 34pp; English.
XX
XX The invention relates to an isolated RecA mutant protein, where the
XX protein is either a single mutant RecA protein comprising a deletion of
XX amino acid residues from the carboxyl terminus, or a double mutant RecA
XX protein comprising a deletion of amino acid residues from the carboxyl

terminus and an amino acid change from a glutamate to a basic amino acid.
The invention also relates to a polynucleotide sequence that encodes the
Reca mutant protein, a method of catalyzing an in vitro homologous DNA
pairing and DNA strand exchange reactions comprising providing an amount
of the RecA mutant protein and a method of increasing recombination
efficiency of homologous DNA pairing and DNA strand exchange reactions in
a cell comprising supplying to the cell an amount of the RecA mutant
protein. The RecA mutant protein comprises an enhanced capacity to
displace a DNA binding protein as compared to wild-type RecA. The DNA
binding protein is the ssDNA binding protein from *Escherichia coli* (SSB).
It also comprises enhanced binding to DNA during a DNA strand exchange
reaction as compared to wild-type RecA. The composition and methods are
useful for catalyzing homologous DNA pairing and DNA strand exchange
reactions in an in vitro or in vivo environment. These may be used in
promoting in vitro alterations of genes to permit the rapid construction
of desired gene mutants for industrial and pharmaceutical purposes. The
mutant proteins may also be used in targeting the site-specific cleavage
of small and large DNAs, or as a basis for the design and construction of
tiny electronic circuits based on DNA. This sequence represents a RecA
mutant protein of the invention.

XX SQ Sequence 335 AA;
Query Match 100.0%; Score 1678; DB 8; Length 335;
Best Local Similarity 100.0%; Pred. No. 2.8e-155;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVKTISTGSLDIALGAGGLPMGR 60
DB 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVKTISTGSLDIALGAGGLPMGR 60
QY 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDINDLLCSQPD 120
DB 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDINDLLCSQPD 120
QY 121 TGEQALEICDALARGAGVDVIVDVSVAALTTPKAEIEGEGSHMGLAARMMSQAMRKLAG 180
DB 121 TGEQALEICDALARGAGVDVIVDVSVAALTTPKAEIEGEGSHMGLAARMMSQAMRKLAG 180
QY 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
DB 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
QY 241 ETRVVKVKNKIAAPFKQAEFOILYEGEINFGYELVDLGVKEKLEKAGAWTSYKGEKIGQ 300
DB 241 ETRVVKVKNKIAAPFKQAEFOILYEGEINFGYELVDLGVKEKLEKAGAWTSYKGEKIGQ 300
QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
DB 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335

RESULT 2

ADRI16428
ID ADRI16428 standard; protein; 335 AA.
AC ADRI16428;
XX 04-NOV-2004 (first entry)
DT
XX E. coli RecA mutant protein #1.
DE
XX RecA; DNA binding protein; ssDNA binding protein; SSB;
KW DNA strand exchange reaction; site-specific cleavage; mutant; muten.
XX
OS *Escherichia coli*.
XX
PN US2004157248-A1.
XX
PD 12-AUG-2004.
XX
PF 11-DEC-2003; 2003US-00733782.
XX

PR 12-DEC-2002; 2002US-0432758P.

XX (COXM/) COX M M.
PA (LUSE/) LUSSETTI S L.
PA (EGL/) EGLER A L.
XX (HARU/) HARUTA N.
PI Cox MM, Lusettti SL, Eggler AL, Haruta N;
XX WPI; 2004-580265/56.
DR N-PSDB; ADR16429.

XX New RecA mutant proteins comprising a single mutation or a double
mutation, useful for catalyzing homologous DNA pairing and DNA strand
exchange reactions in an in vitro or in vivo environment.

XX Claim 2; SEQ ID NO 1; 34pp; English.

CC The invention relates to an isolated RecA mutant protein, where the
protein is either a single mutant RecA protein comprising a deletion of
amino acid residues from the carboxyl terminus, or a double mutant RecA
protein comprising a deletion of amino acid residues from the carboxyl
terminus and an amino acid change from a glutamate to a basic amino acid.
The invention also relates to a polynucleotide sequence that encodes the
Reca mutant protein, a method of catalyzing an in vitro homologous DNA
pairing and DNA strand exchange reactions comprising providing an amount
of the RecA mutant protein and a method of increasing recombination
efficiency of homologous DNA pairing and DNA strand exchange reactions in
a cell comprising supplying to the cell an amount of the RecA mutant
protein. The RecA mutant protein comprises an enhanced capacity to
displace a DNA binding protein as compared to wild-type RecA. The DNA
binding protein is the ssDNA binding protein from *Escherichia coli* (SSB).
It also comprises enhanced binding to DNA during a DNA strand exchange
reaction as compared to wild-type RecA. The composition and methods are
useful for catalyzing homologous DNA pairing and DNA strand exchange
reactions in an in vitro or in vivo environment. These may be used in
promoting in vitro alterations of genes to permit the rapid construction
of desired gene mutants for industrial and pharmaceutical purposes. The
mutant proteins may also be used in targeting the site-specific cleavage
of small and large DNAs, or as a basis for the design and construction of
tiny electronic circuits based on DNA. This sequence represents a RecA
mutant protein of the invention.

XX SQ Sequence 335 AA;

Query Match 99.8%; Score 1674; DB 8; Length 335;
Best Local Similarity 99.7%; Pred. No. 7e-155;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVKTISTGSLDIALGAGGLPMGR 60
DB 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVKTISTGSLDIALGAGGLPMGR 60
QY 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDINDLLCSQPD 120
DB 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDINDLLCSQPD 120
QY 121 TGEQALEICDALARGAGVDVIVDVSVAALTTPKAEIEGEGSHMGLAARMMSQAMRKLAG 180
DB 121 TGEQALEICDALARGAGVDVIVDVSVAALTTPKAEIEGEGSHMGLAARMMSQAMRKLAG 180
QY 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
DB 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
QY 241 ETRVVKVKNKIAAPFKQAEFOILYEGEINFGYELVDLGVKEKLEKAGAWTSYKGEKIGQ 300
DB 241 ETRVVKVKNKIAAPFKQAEFOILYEGEINFGYELVDLGVKEKLEKAGAWTSYKGEKIGQ 300
QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
DB 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335

RESULT 3
ADRI16432
ID ADRI16432 standard; protein; 352 AA.
XX
AC ADRI16432;
XX
DT 04-NOV-2004 (first entry)
XX
DE E. coli wild-type RecA protein.
XX
KW RecA; DNA binding protein; ssDNA binding protein; SSB;
KW DNA strand exchange reaction; site-specific cleavage.
XX
OS Escherichia coli.
XX
PN US2004157248-A1.
XX
PD 12-AUG-2004.
XX
PF 11-DEC-2003; 2003US-00733782.
XX
PR 12-DEC-2002; 2002US-0432758P.
XX
PA (COXM//) COX M M.
PA (LUSE//) LUSETTI S L.
PA (EGGL//) EGGLE A L.
PA (HARU//) HARUTA N.
XX
PI Cox MM, Lusetti SL, Eggler AL, Haruta N;
XX
DR WPI; 2004-580265/56.
XX
XX
XX New RecA mutant proteins comprising a single mutation or a double
PT mutation, useful for catalyzing homologous DNA pairing and DNA strand
PT exchange reactions in an in vitro or in vivo environment.
XX
XX
XX Disclosure; SEQ ID NO 5; 34pp; English.
XX
XX The invention relates to an isolated RecA mutant protein, where the
CC protein is either a single mutant RecA protein comprising a deletion of
CC amino acid residues from the carboxyl terminus, or a double mutant RecA
CC protein comprising a deletion of amino acid residues from the carboxyl
CC terminus and an amino acid change from a glutamate to a basic amino acid.
CC The invention also relates to a polynucleotide sequence that encodes the
CC RecA mutant protein, a method of catalyzing an in vitro homologous DNA
CC pairing and DNA strand exchange reactions comprising providing an amount
CC of the RecA mutant protein and a method of increasing recombination
CC efficiency of homologous DNA pairing and DNA strand exchange reactions in
CC a cell comprising supplying to the cell an amount of the RecA mutant
CC protein. The RecA mutant protein comprises an enhanced capacity to
CC displace a DNA binding protein as compared to wild-type RecA. The DNA
CC binding protein is the ssDNA binding protein from Escherichia coli (SSB).
CC It also comprises enhanced binding to DNA during a DNA strand exchange
CC reaction as compared to wild-type RecA. The composition and methods are
CC useful for catalyzing homologous DNA pairing and DNA strand exchange
CC reactions in an in vitro or in vivo environment. These may be used in
CC promoting in vitro alterations of genes to permit the rapid construction
CC of desired gene mutants for industrial and pharmaceutical purposes. The
CC mutant proteins may also be used in targeting the site-specific cleavage
CC of small and large DNAs, or as a basis for the design and construction of
CC tiny electronic circuits based on DNA. This sequence represents the wild-
CC type RecA protein of the invention.
XX
XX
XX Sequence 352 AA;
XX
XX Query Match 99.8%; Score 1674; DB 8; Length 352;
XX Best Local Similarity 99.7%; Pred. NO. 7.5e-155;
XX Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGDRSDMDVETISTGSLDLIALGAGGLPMGR 60
XX
XX 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGDRSDMDVETISTGSLDLIALGAGGLPMGR 60
XX

QY 61 IVEIYGPESSGKTTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVVDINLLCSQPD 120
DB 61 IVEIYGPESSGKTTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVVDINLLCSQPD 120
QY 121 TGEQALEICDALARSGAVDVIWVDSVAALTPKAEIETGEIGDSHMGGLAARMMSQAMRKLAG 180
DB 121 TGEQALEICDALARSGAVDVIWVDSVAALTPKAEIETGEIGDSHMGGLAARMMSQAMRKLAG 180
QY 181 NLKOSNTLLIFINQIRMKIGVMFGNPETTTGNGALKFYASVRLDIRRGAVKEGNNVGS 240
DB 181 NLKOSNTLLIFINQIRMKIGVMFGNPETTTGNGALKFYASVRLDIRRGAVKEGNNVGS 240
QY 241 ETRVKVKNKIAAPFKQAEFQILYCEGINFYGELVDLGVKBLIEKAGAWSYKGEKIQG 300
DB 241 ETRVKVKNKIAAPFKQAEFQILYCEGINFYGELVDLGVKBLIEKAGAWSYKGEKIQG 300
QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 335
DB 301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 335
RESULT 4
ABUL5395
ID ABUL5395 standard; protein; 353 AA.
XX
AC ABUL5395;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #922.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Escherichia coli.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
XX
XX N-ESDB; ACA19265.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 43319; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of


```

XX 16-MAY-2000 (first entry)
XX Amino acid sequence of a hyperrecombinogenic recA protein clone 4.
XX recA; hyperrecombinogenic variant; male gamete; female gamete;
XX taxol; ethanol production; ethanol tolerance; metabolite; taxol;
XX cyclosporin A; erythromycin; meiosis.
XX Synthetic.
XX Escherichia coli.
XX WO200004190-A1.
XX 27-JAN-2000.
XX 15-JUL-1999; 99WO-US015972.
XX 15-JUL-1998; 98US-00116188.
XX (MAXY-) MAXYGEN INC.
XX Del Cardayre S, Tobin M, Stemmer WPC, Ness JE, Minshull J;
XX Patten PA, Subramanian V, Castle LA, Krebber CM, Bass S, Zhang Y;
XX Cox T, Huisman G, Yuan L, Affholter JA;
XX WPI; 2000-182446/16.
XX N-PSDB; NAZ60611.
XX Evolution of whole cells and organism by iterative cycles of
XX recombination and selection and screening for acquisition of desired
XX properties.
XX Example 1; Fig 13; 197pp; English.
XX The present sequence represents a hyperrecombinogenic recA protein. The
XX wild type recA protein (AAV68827) was used to produce hyperrecombinogenic
XX variants, using the method of the invention. The specification describes
XX a method for producing a library of diverse multicellular organisms using
XX pools of male and female gametes. At least one of the male pool or female
XX pools comprises a number of different gametes derived from different
XX strains of a species or of a different species. The viable organisms
XX produced from the fertilized gametes are repeatedly crossed to produce a
XX library of diverse organisms, which are selected for a desired trait or
XX property. The methods can be used to evolve cells to acquire a desired
XX property such as heat tolerance, ethanol production or tolerance, acid,
XX improved production and maintenance of enzyme cofactors or NAD(P)H and
XX improved glucose transport. The desired property may be expression of a
XX protein or primary or secondary metabolite. Alternatively the desired
XX property is secretion of a protein or secondary metabolite, chosen from
XX taxol, cyclosporin A and erythromycin. The desired property may be a
XX capacity for meiosis or compatibility to form a heterokaryon with another
XX strain
XX SQ Sequence 358 AA;
Query Match 99.8%; Score 1674; DB 3; Length 358;
Best Local Similarity 99.7%; Pred. No. 7.7e-155;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALQIEKQFGKGSIMRLGEDRSMVDVTISTGSLDLIALGAGLPMGR 60
DB 7 AIDENKQKALAAALQIEKQFGKGSIMRLGEDRSMVDVTISTGSLDLIALGAGLPMGR 66
QY 61 IVEYGPSSGKTTLTQVIAAAREGKTCAFDAEHALDPYARKLGVDINDLLCSQPD 120
DB 67 IVEYGPSSGKTTLTQVIAAAREGKTCAFDAEHALDPYARKLGVDINDLLCSQPD 126
QY 121 TGEQALIEICALARSGAVDVIVDSVAALTPKABIEGIGDSHMLAARWMSQAMRKLKAG 180
DB 127 TGEQALIEICALARSGAVDVIVDSVAALTPKABIEGIGDSHMLAARWMSQAMRKLKAG 186
QY 181 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGVS 240

```

```

DB 187 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGVS 246
QY 241 ETRVKKVKNKIAAPFKQAEFOILYGEINFGYGLVDLGVKKEKLEKAGAWSYKGEKIGQ 300
DB 247 ETRVKKVKNKIAAPFKQAEFOILYGEINFGYGLVDLGVKKEKLEKAGAWSYKGEKIGQ 306
QY 301 GKANATAWLKDNPETAKEIEKKVRELLLSNPNSTP 335
DB 307 GKANATAWLKDNPETAKEIEKKVRELLLSNPNSTP 341
RESULT 7
ADQ87824
ID ADQ87824 standard; protein; 358 AA.
XX AC ADQ87824;
XX DT 09-SEP-2004 (first entry)
XX DE E. coli recA protein clone #3.
XX KW recA; protoplast formation; gamete; heat tolerance; ethanol production;
XX ethanol tolerance; enzyme cofactor; NAD(P)H; glucose transport; meiosis;
XX heterokaryon; taxol; cyclosporin A; erythromycin; viral infection;
XX plant genome shuffling; micropore manipulation; reiterative pooling.
XX OS Escherichia coli.
XX FN AU2004200501-A1.
XX PD 04-MAR-2004.
XX PF 09-FEB-2004; 2004AU-00200501.
XX PR 09-FEB-2004; 2004AU-00200501.
XX PA (MAXY-) MAXYGEN INC.
XX PI Yuan L, Cox T, Bass S, Castle LA, Patten PA, Ness JE, Tobin M;
XX Affholter JA, Huisman G, Zhang Y, Krebber CM, Subramanian V;
XX Minshull J, Stemmer WPC, Del Cardayre S;
XX WPI; 2004-507924/49.
XX N-PSDB; ADQ87817.
XX Evolving cells to acquire a desired property, by forming protoplasts of
XX different cells, fusing protoplasts to form hybrid protoplasts (HP),
XX producing regenerated cells, forming additional HP and producing
XX additional regenerated cells.
XX PS Disclosure; Fig 13; 196pp; English.
XX The invention relates to a method of evolving cells to acquire a desired
XX property by forming protoplasts of different cells, fusing the
XX protoplasts to form hybrid protoplasts, incubating the hybrid protoplasts
XX to produce regenerated cells, repeatedly forming protoplasts from
XX regenerated cells, fusing protoplasts to form hybrid protoplasts in which
XX genomes from protoplasts recombine to form additional hybrid genomes and
XX incubating additional hybrid protoplasts for producing additional
XX regenerated cells. The invention also relates to a method of producing a
XX library of diverse multicellular organisms involving providing a pool of
XX male gametes and a pool of female gametes, where one of the male pool or
XX the female pool comprises several different gametes derived from
XX different strains of a species or different species and the male gametes
XX fertilise the female gametes, permitting at least a portion of the
XX resulting fertilised gametes to grow into reproductively viable
XX organisms, repeatedly crossing the reproductively viable organisms to
XX produce a library of diverse organisms and selecting the library for a
XX desired trait or property. The methods are useful for evolving cells to
XX acquire desired properties such as heat tolerance, ethanol production,
XX ethanol tolerance, improved production and maintenance of enzyme
XX cofactors, improved production and maintenance of NAD(P)H and improved

```

CC glucose transport. The desired property is the expression of a protein,
CC primary metabolite or secondary metabolite, the secretion of a protein or
CC secondary metabolite, capacity for meiosis or compatibility to form a
CC heterokaryon with another strain. The secondary metabolite is chosen from
CC taxol, cyclosporin A and erythromycin. The method is also useful for
CC acquisition of desired properties such as enhanced recombinationity,
CC gene copy number or gene reductivity and capacity for expression and/or
CC secretion of proteins or secondary metabolites. The method is further
CC useful for predicting efficacy of a drug in treating viral infection,
CC plant genome shuffling, micropore manipulation, producing transgenic
CC animals, improvement of overexpressed genes for a desired phenotype and
CC reiterative pooling and breeding of higher organisms. The cells evolved
CC by the method are useful in molecular genetics. This sequence represents
CC an E. coli recA protein clone used in the method of the invention.

XX Sequence 358 AA;

Query Match 99.8%; Score 1674; DB 8; Length 358;
Best Local Similarity 99.7%; Pred. No. 7.7e-155;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPSSCKTTLTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDDINLLCSQPD 120
Db 67 IVEIYGPSSCKTTLTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDDINLLCSQPD 126
Qy 121 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGIGSHMGLAARMMSQMRKLAG 180
Db 127 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGIGSHMGLAARMMSQMRKLAG 186
Qy 181 NLKQSNLTLLIFINQIRMKIGVFNPGPPTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
Db 187 NLKQSNLTLLIFINQIRMKIGVFNPGPPTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 246
Qy 241 ETRVVKVKNKIAAPFKQAEFQILYEGINFGYELVDLGVKEKLIIEKAGAWYSYKGEKIGQ 300
Db 247 ETRVVKVKNKIAAPFKQAEFQILYEGINFGYELVDLGVKEKLIIEKAGAWYSYKGEKIGQ 306
Qy 301 GKANATAWLKONPETAKIEKKVRELLLSNPNSTP 335
Db 307 GKANATAWLKONPETAKIEKKVRELLLSNPNSTP 341

RESULT 8
ADQ87822

ID ADQ87822 standard; protein; 358 AA.

XX ADQ87822;

XX 09-SEP-2004 (first entry)

XX E. coli recA protein clone #1.

XX RecA; protoplast formation; gamete; heat tolerance; ethanol production;
KW ethanol tolerance; enzyme cofactor; NAD(P)H; glucose transport; meiosis;
KW heterokaryon; taxol; cyclosporin A; erythromycin; viral infection;
KW plant genome shuffling; micropore manipulation; reiterative pooling.

XX Escherichia coli.

XX AU2004200501-A1.

XX 04-MAR-2004.

XX 09-FEB-2004; 2004AU-00200501.

XX 09-FEB-2004; 2004AU-00200501.

XX (MAXY-) MAXYGEN INC.

PI Yuan L, Cox T, Baas S, Castle LA, Patten PA, Ness JB, Tobin M;
PI Affholter JA, Huismann G, Zhang Y, Kriebber CM, Subramanian V;
XX Minshull J, Stemmer WPC, Del Cardayre S;

DR WPI; 2004-507924/49.

DR N-PSDB; ADQ87815.

XX Evolving cells to acquire a desired property, by forming protoplasts of
PT different cells, fusing protoplasts to form hybrid protoplasts (HP),
PT producing regenerated cells, forming additional HP and producing
XX additional regenerated cells.

PS Disclosure; Fig 13; 196pp; English.

XX The invention relates to a method of evolving cells to acquire a desired
CC property by forming protoplasts of different cells, fusing the
CC protoplasts to form hybrid protoplasts, incubating the hybrid protoplasts
CC to produce regenerated cells, repeatedly forming protoplasts from
CC regenerated cells, fusing protoplasts to form hybrid protoplasts in which
CC genomes from protoplasts recombine to form additional hybrid genomes and
CC incubating additional hybrid protoplasts for producing additional
CC regenerated cells. The invention also relates to a method of producing a
CC library of diverse multicellular organisms involving providing a pool of
CC male gametes and a pool of female gametes, where one of the male pool or
CC the female pool comprises several different gametes derived from
CC different strains of a species or different species and the male gametes
CC fertilise the female gametes, permitting at least a portion of the
CC resulting fertilised gametes to grow into reproductively viable
CC organisms, repeatedly crossing the reproductively viable organisms to
CC produce a library of diverse organisms and selecting the library for a
CC desired trait or property. The methods are useful for evolving cells to
CC acquire desired properties such as heat tolerance, ethanol production,
CC ethanol tolerance, improved production and maintenance of enzyme
CC cofactors, improved production and maintenance of NAD(P)H and improved
CC glucose transport. The desired property is the expression of a protein, or
CC primary metabolite or secondary metabolite, the secretion of a protein or
CC secondary metabolite, capacity for meiosis or compatibility to form a
CC heterokaryon with another strain. The secondary metabolite is chosen from
CC taxol, cyclosporin A and erythromycin. The method is also useful for
CC acquisition of desired properties such as enhanced recombinationity,
CC gene copy number or gene reductivity and capacity for expression and/or
CC secretion of proteins or secondary metabolites. The method is further
CC useful for predicting efficacy of a drug in treating viral infection,
CC plant genome shuffling, micropore manipulation, producing transgenic
CC animals, improvement of overexpressed genes for a desired phenotype and
CC reiterative pooling and breeding of higher organisms. The cells evolved
CC by the method are useful in molecular genetics. This sequence represents
CC an E. coli recA protein clone used in the method of the invention.

XX Sequence 358 AA;

Query Match 99.8%; Score 1674; DB 8; Length 358;

Best Local Similarity 99.7%; Pred. No. 7.7e-155;

Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 60

Db 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPSSCKTTLTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDDINLLCSQPD 120

Db 67 IVEIYGPSSCKTTLTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDDINLLCSQPD 126

Qy 121 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGIGSHMGLAARMMSQMRKLAG 180

Db 127 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGIGSHMGLAARMMSQMRKLAG 186

Qy 181 NLKQSNLTLLIFINQIRMKIGVFNPGPPTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240

Db 187 NLKQSNLTLLIFINQIRMKIGVFNPGPPTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 246

Qy 241 ETRVVKVKNKIAAPFKQAEFQILYEGINFGYELVDLGVKEKLIIEKAGAWYSYKGEKIGQ 300

Db 247 ETRVVKVKNKTAAPFKQAEFOILYGEINFGYELVDLGVKEKLEKAGAWSYKGEKIGQ 306

Qy 301 GKANATWLKNDPETAKEIEKKVRELLLSNPSTP 335

Db 307 GKANATWLKNDPETAKEIEKKVRELLLSNPSTP 341

RESULT 9

ADQ87821

ID ADQ87821 standard; protein; 358 AA.

XX

AC ADQ87821;

DT 09-SEP-2004 (first entry)

XX

DE E. coli recA protein.

XX

KW RecA; protoplast formation; gamete; heat tolerance; ethanol production;

KW ethanol tolerance; enzyme cofactor; NAD(P)H; glucose transport; meiosis;

KW heterokaryon; taxol; cyclosporin A; erythromycin; viral infection;

KW plant genome shuffling; micropore manipulation; reiterative pooling.

OS Escherichia coli.

XX

FN AU2004200501-A1.

XX

PD 04-MAR-2004.

XX

PF 09-FEB-2004; 2004AU-00200501.

XX

PR 09-FEB-2004; 2004AU-00200501.

XX

PA (MAXY-) MAXYGEN INC.

XX

PI Yuan L, Cox T, Bass S, Castle LA, Patten PA, Ness JE, Tobin M;

PI Affolter JA, Huismann G, Zhang Y, Krebber CW, Subramanian V;

PI Minshull J, Stemmer WPC, Del Cardayre S;

XX

DR WPI; 2004-507924/49.

DR N-PSDB; ADQ87814.

XX

PT Evolving cells to acquire a desired property, by forming protoplasts of

PT different cells, fusing protoplasts to form hybrid protoplasts (HP),

PT producing regenerated cells, forming additional HP and producing

PT additional regenerated cells.

XX

PS Disclosure; Fig 13; 196pp; English.

XX

CC The invention relates to a method of evolving cells to acquire a desired

CC property by forming protoplasts of different cells, fusing the

CC protoplasts to form hybrid protoplasts, incubating the hybrid protoplasts

CC to produce regenerated cells, repeatedly forming protoplasts from

CC regenerated cells, fusing protoplasts to form hybrid protoplasts in which

CC genomes from protoplasts recombine to form additional hybrid genomes and

CC incubating additional hybrid protoplasts for producing additional

CC regenerated cells. The invention also relates to a method of producing a

CC library of diverse multicellular organisms involving providing a pool of

CC male gametes and a pool of female gametes, where one of the male pool or

CC the female pool comprises several different gametes derived from

CC different strains of a species or different species and the male gametes

CC fertilise the female gametes, permitting at least a portion of the

CC resulting fertilised gametes to grow into reproductively viable

CC organisms, repeatedly crossing the reproductively viable organisms to

CC produce a library of diverse organisms and selecting the library for a

CC desired trait or property. The methods are useful for evolving cells to

CC acquire desired properties such as heat tolerance, ethanol production,

CC ethanol tolerance, improved production and maintenance of enzyme

CC cofactors, improved production and maintenance of NAD(P)H and improved

CC glucose transport. The desired property is the expression of a protein,

CC primary metabolite or secondary metabolite, the secretion of a protein or

CC secondary metabolite, capacity for meiosis or compatibility to form a

CC heterokaryon with another strain. The secondary metabolite is chosen from

CC taxol, cyclosporin A and erythromycin. The method is also useful for

CC acquisition of desired properties such as enhanced recombinationity,

CC gene copy number or gene reductivity and capacity for expression and/or

CC secretion of proteins or secondary metabolites. The method is further

CC useful for predicting efficacy of a drug in treating viral infection,

CC plant genome shuffling, micropore manipulation, producing transgenic

CC animals, improvement of overexpressed genes for a desired phenotype and

CC reiterative pooling and breeding of higher organisms. The cells evolved

CC by the method are useful in molecular genetics. This sequence represents

CC the E. coli recA protein used in the method of the invention.

XX

SQ Sequence 358 AA;

Query Match 99.8%; Score 1674; DB 8; Length 358;

Best Local Similarity 99.7%; Pred. No. 7.7e-155;

Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGSDRSDVKTISTGSLSLDIALGAGGLPMGR 60

Db 7 AIDENKQKALAAALGOIEKQFGKGSIMRLGSDRSDVKTISTGSLSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPSSGKTTLTLOVIAAAOREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120

Db 67 IVEIYGPSSGKTTLTLOVIAAAOREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 126

Qy 121 TGEQALEICDALARSGAVDVIVVDSVAALTTPKABIEGEIGDSHMLAARMMSQAMRKLKAG 180

Db 127 TGEQALEICDALARSGAVDVIVVDSVAALTTPKABIEGEIGDSHMLAARMMSQAMRKLKAG 186

Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240

Db 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 246

Qy 241 ETRVVKVKNKTAAPFKQAEFOILYGEINFGYELVDLGVKEKLEKAGAWSYKGEKIGQ 300

Db 247 ETRVVKVKNKTAAPFKQAEFOILYGEINFGYELVDLGVKEKLEKAGAWSYKGEKIGQ 306

Qy 301 GKANATWLKNDPETAKEIEKKVRELLLSNPSTP 335

Db 307 GKANATWLKNDPETAKEIEKKVRELLLSNPSTP 341

RESULT 10

AA568831

ID AA568831 standard; protein; 358 AA.

XX

AC AA568831;

DT 16-MAY-2000 (first entry)

XX

DE Amino acid sequence of a hyperrecombinogenic recA protein clone 6.

XX

KW recA; hyperrecombinogenic variant; male gamete; female gamete;

KW heat tolerance; ethanol production; ethanol tolerance; taxol;

KW cyclosporin A; erythromycin; meiosis.

XX

OS Synthetic.

OS Escherichia coli.

XX

PH Key Location/Qualifiers

FT Misc-difference 156

FT /note= "encoded by TCG"

XX

PN WO200004190-A1.

XX

PD 27-JAN-2000.

XX

PF 15-JUL-1999; 99WO-US015972.

XX

PR 15-JUL-1998; 98US-00116188.

XX

PA (MAXY-) MAXYGEN INC.

XX

PI Del Cardayre S, Tobin M, Stemmer WPC, Ness JE, Minshull J;

PI Patten PA, Subramanian V, Castle LA, Krebber CM, Bass S, Zhang Y;
 PI Cox T, Huisman G, Yuan L, Affholter JA;
 DR WPI; 2000-182446/16.
 DR N-PSDB; AAZ60613.
 XX Evolution of whole cells and organism by iterative cycles of
 PT recombination and selection and screening for acquisition of desired
 PT properties.
 XX Example 1; Fig 13; 197pp; English.
 XX The present sequence represents a hyperrecombinogenic recA protein. The
 CC wild type recA protein (AAV68827) was used to produce hyperrecombinogenic
 CC variants, using the method of the invention. The specification describes
 CC a method for producing a library of diverse multicellular organisms using
 CC pools of male and female gametes. At least one of the male pool or female
 CC pools comprises a number of different gametes derived from different
 CC strains of a species or of a different species. The viable organisms
 CC produced from the fertilized gametes are repeatedly crossed to produce a
 CC library of diverse organisms, which are selected for a desired trait or
 CC property. The methods can be used to evolve cells to acquire a desired
 CC property such as heat tolerance, ethanol production or tolerance, acid,
 CC improved production and maintenance of enzyme cofactors or NAD(P)H and
 CC improved glucose transport. The desired property may be expression of a
 CC protein or primary or secondary metabolite. Alternatively the desired
 CC property is secretion of a protein or secondary metabolite, chosen from
 CC taxol, cyclosporin A and erythromycin. The desired property may be a
 CC capacity for meiosis or compatibility to form a heterokaryon with another
 CC strain
 XX Sequence 358 AA;

Query Match 99.7%; Score 1673; DB 3; Length 358;
 Best Local Similarity 99.4%; Pred. No. 9.7e-155;
 Matches 333; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVKTISTGSLSLDIALGAGGLPMGR 60
 DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVKTISTGSLSLDIALGAGGLPMGR 66
 QY 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDDNLLCSQPD 120
 DB 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDDNLLCSQPD 126
 QY 121 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 180
 DB 127 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 186
 QY 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVYGS 240
 DB 187 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVYGS 246
 QY 241 ETRVKVKNKIAAPFKQAEFOILYEGEINFGELVDLGVKEKLEKAGAWTSYKGEKIGQ 300
 DB 247 ETRVKVKNKIAAPFKQAEFOILYEGEINFGELVDLGVKEKLEKAGAWTSYKGEKIGQ 306
 QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 335
 DB 307 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 341

RESULT 11
 AAW64213
 ID AAW64213 standard; protein; 358 AA.
 XX AAW64213;
 XX 28-APR-1999 (first entry)
 DT New minshall recA protein.
 DE Escherichia coli; recA protein; hyperrecombinogenic variant; Minshall;
 KW

KW recursive sequence recombination; evolution.
 XX Escherichia coli.
 OS WO9831837-A1.
 XX 23-JUL-1998.
 PD 16-JAN-1998; 98WO-US000852.
 PF 17-JAN-1997; 97US-0035054P.
 PR (MAXY-) MAXYGEN INC.
 PA Delcardayre SB, Tobin MB, Stemmer WPC, Ness JE, Minshull J;
 PI Patten P;
 XX WPI; 1998-427565/60.
 DR N-PSDB; AAV44285.
 XX Evolution of whole cells and organisms by recursive DNA sequence
 PT recombination in cells to evolve cells having acquired desired function,
 PT useful in methods for predicting the efficacy of a drug in treating viral
 PT or pathogenic infections.
 XX Example 1; Fig 13; 125pp; English.
 PS The present invention provides methods employing iterative cycles of
 CC recombination and selection/screening for evolution of whole cells and
 CC organisms toward acquisition of desired properties, e.g. enhanced
 CC recombinogenicity, genome copy number, and capacity for expression and/or
 CC secretion of proteins and secondary metabolites. The present sequence
 CC represents a wild-type recA protein (designated new Minshall), from an
 CC example of the present invention
 XX Sequence 358 AA;

Query Match 99.5%; Score 1670; DB 2; Length 358;
 Best Local Similarity 99.4%; Pred. No. 1.9e-154;
 Matches 333; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVKTISTGSLSLDIALGAGGLPMGR 60
 DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVKTISTGSLSLDIALGAGGLPMGR 66
 QY 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDDNLLCSQPD 120
 DB 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDDNLLCSQPD 126
 QY 121 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 180
 DB 127 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 186
 QY 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVYGS 240
 DB 187 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVYGS 246
 QY 241 ETRVKVKNKIAAPFKQAEFOILYEGEINFGELVDLGVKEKLEKAGAWTSYKGEKIGQ 300
 DB 247 ETRVKVKNKIAAPFKQAEFOILYEGEINFGELVDLGVKEKLEKAGAWTSYKGEKIGQ 306
 QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 335
 DB 307 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 341

RESULT 12
 AAY68828
 ID AAY68828 standard; protein; 358 AA.
 XX AAY68828;
 XX 16-MAY-2000 (first entry)
 DT

XX
DE Amino acid sequence of a hyperrecombinogenic recA protein clone 2.
XX
KW recA; hyperrecombinogenic variant; male gamete; female gamete;
KW heat tolerance; ethanol production; ethanol tolerance; metabolite; taxol;
KW cyclosporin A; erythromycin; meiosis.

Qy	241	ETRVKVVKNKIAAPFKQAEFQQLLYGEGINFYVYVLDLGVKKEKLIKAGAWYSYKGEKIQ	300
Ds	247	ETRVKVVKNKIAAPFKQAEFQQLLYGEGINFYVYVLDLGVKKEKLIKAGAWYSYKGEKIQ	306
Qy	301	GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP	335
Ds	307	GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP	341
RESULT 13			
ID	ADQ87823		
XX	ADQ87823 standard; protein; 358 AA.		
AC	ADQ87823;		
DT	09-SEP-2004 (first entry)		
XX	E. coli recA protein clone #2.		
DE			
XX	RecA; protoplast formation; gamete; heat tolerance; ethanol production;		
KW	ethanol tolerance; enzyme cofactor; NAD(P)H; glucose transport; meiosis;		
KW	heterokaryon; taxol; cyclosporin A; erythromycin; viral infection;		
KW	plant genome shuffling; micropore manipulation; reiterative pooling.		
XX			
OS	Escherichia coli.		
XX			
FN	AU2004200501-A1.		
PD	04-MAR-2004.		
XX			
PF	09-FEB-2004; 2004AU-00200501.		
XX			
FR	09-FEB-2004; 2004AU-00200501.		
XX	(MAXY-) MAXYGEN INC.		
PI	Yuan L, Cox T, Baas S, Castle LA, Patten PA, Ness JE, Tobin M;		
PI	Affholter JA, Huisman G, Zhang Y, Krebber CW, Subramanian V;		
PI	Minshull J, Stember WPC, Del Cardayre S;		
XX			
DR	WPI; 2004-507924/49.		
DR	N-PSDB; ADQ87816.		
XX			
PT	Evolving cells to acquire a desired property, by forming protoplasts of		
PT	different cells, fusing protoplasts to form hybrid protoplasts (HP),		
PT	producing regenerated cells, forming additional HP and producing		
PT	additional regenerated cells.		
PS	Disclosure; Fig 13; 196pp; English.		
XX			
CC	The invention relates to a method of evolving cells to acquire a desired		
CC	property by forming protoplasts of different cells, fusing the		
CC	protoplasts to form hybrid protoplasts, incubating the hybrid protoplasts		
CC	to produce regenerated cells, repeatedly forming protoplasts from		
CC	regenerated cells, fusing protoplasts to form hybrid protoplasts in which		
CC	genomes from protoplasts recombine to form additional hybrid genomes and		
CC	incubating additional hybrid protoplasts for producing additional		
CC	regenerated cells. The invention also relates to a method of producing a		
CC	library of diverse multicellular organisms involving providing a pool of		
CC	male gametes and a pool of female gametes, where one of the male pool or		
CC	the female pool comprises several different gametes derived from		
CC	different strains of a species or different species and the male gametes		
CC	fertilise the female gametes, permitting at least a portion of the		
CC	resulting fertilised gametes to grow into reproductively viable		
CC	organisms, repeatedly crossing the reproductively viable organisms to		
CC	produce a library of diverse organisms and selecting the library for a		
CC	desired trait or property. The methods are useful for evolving cells to		
CC	acquire desired properties such as heat tolerance, ethanol production,		
CC	ethanol tolerance, improved production and maintenance of enzyme		
CC	cofactors, improved production and maintenance of NAD(P)H and improved		
CC	glucose transport. The desired property is the expression of a protein,		
CC	primary metabolite or secondary metabolite, the secretion of a protein or		

CC secondary metabolite, capacity for meiosis or compatibility to form a
CC heterokaryon with another strain. The secondary metabolite is chosen from
CC taxol, cyclosporin A and erythromycin. The method is also useful for
CC acquisition of desired properties such as enhanced recombination, gene
CC copy number or gene reductivity and capacity for expression and/or
CC secretion of proteins or secondary metabolites. The method is further
CC useful for predicting efficacy of a drug in treating viral infection,
CC plant genome shuffling, micropore manipulation, producing transgenic
CC animals, improvement of overexpressed genes for a desired phenotype and
CC reiterative pooling and breeding of higher organisms. The cells evolved
CC by the method are useful in molecular genetics. This sequence represents
CC an E. coli recA protein clone used in the method of the invention.

XX Sequence 358 AA;

Query Match 99.5%; Score 1669; DB 8; Length 358;

Best Local Similarity 99.1%; Pred. No. 2.4e-154;

Matches 332; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AIDENKQALAAALGQIEKQFGKGSIMRLGDRSDVKTISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQALATAGQIEKQFGKGSIMRLGDRSDVKTISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDDNLLCSQPD 120
Db 67 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDDNLLCSQPD 126
Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTTPKAEIEGIGDSHMGLAARMMSQAMRKLAG 180
Db 127 TGEQALEICDALARSGAVDVIVDSVAALTTPKAEIEGIGDSHMGLAARMMSQAMRKLAG 186
Qy 181 NLKQSNLTLLIFINQIRMKIGWFGNPNPTTTCGNALKFYASVRLDIRRIGAVKEGENVYGS 240
Db 187 NLKQSNLTLLIFINQIRMKIGWFGNPNPTTTCGNALKFYASVRLDIRRIGAVKEGENVYGS 246
Qy 241 ETRVKVKNKIAAPFKQAEFQILYEGEINFYGVLDLGVKEKLIBKAGAWYSYKEKIGQ 300
Db 247 ETRVKVKNKIAAPFKQAEFQILYEGEINFYGVLDLGVKEKLIBKAGAWYSYKEKIGQ 306
Qy 301 GKANATAMKONPETAKIEKKVRELLLSNPNSTP 335
Db 307 GKANATAMKONPETAKIEKKVRELLLSNPNSTP 341

RESULT 14

ADQ87826

ID ADQ87826 standard; protein; 358 AA.

XX AC ADQ87826;

XX DT 09-SEP-2004 (first entry)

XX DE E. coli recA protein clone #5.

XX KW RecA; protoplast formation; gamete; heat tolerance; ethanol production;
KW ethanol tolerance; enzyme cofactor; NAD(P)H; Glucose transport; meiosis;
KW heterokaryon; taxol; cyclosporin A; erythromycin; viral infection;
KW plant genome shuffling; micropore manipulation; reiterative pooling.

XX OS Escherichia coli.

XX PN AU2004200501-A1.

XX PD 04-MAR-2004.

XX PF 09-FEB-2004; 2004AU-00200501.

XX PR 09-FEB-2004; 2004AU-00200501.

XX PA (MAXY-) MAXYGEN INC.

XX PI Yuan L, Cox T, Baas S, Caetle LA, Patten PA, Ness JE, Tobin M;

PI Affholter JA, Huisman G, Zhang Y, Krebber CM, Subramanian V;

PI

XX Minshull J, Stemmer WPC, Del Cardayre S;

DR WPI; 2004-507924/49.

XX N-PSDB; ADQ87819.

XX Evolving cells to acquire a desired property, by forming protoplasts of
XX different cells, fusing protoplasts to form hybrid protoplasts (HP),
XX producing regenerated cells, forming additional HP and producing
XX additional regenerated cells.

XX Disclosure; Fig 13; 196pp; English.

XX The invention relates to a method of evolving cells to acquire a desired
XX property by forming protoplasts of different cells, fusing the
XX protoplasts to form hybrid protoplasts, incubating the hybrid protoplasts
XX to produce regenerated cells, repeatedly forming protoplasts from
XX regenerated cells, fusing protoplasts to form hybrid protoplasts in which
XX genomes from protoplasts recombine to form additional hybrid genomes and
XX incubating additional hybrid protoplasts for producing additional
XX regenerated cells. The invention also relates to a method of producing a
XX library of diverse multicellular organisms involving providing a pool of
XX male gametes and a pool of female gametes, where one of the male pool or
XX the female pool comprises several different gametes derived from
XX different strains of a species or different species and the male gametes
XX fertilise the female gametes, permitting at least a portion of the
XX resulting fertilised gametes to grow into reproductively viable
XX organisms, repeatedly crossing the reproductively viable organisms to
XX produce a library of diverse organisms and selecting the library for a
XX desired trait or properties. The methods are useful for evolving cells to
XX acquire desired properties such as heat tolerance, ethanol production,
XX ethanol tolerance, improved production and maintenance of enzyme
XX cofactors, improved production and maintenance of NAD(P)H and improved
XX glucose transport. The desired property is the expression of a protein,
XX primary metabolite or secondary metabolite, the secretion of a protein or
XX heterokaryon with another strain. The secondary metabolite is chosen from
XX taxol, cyclosporin A and erythromycin. The method is also useful for
XX acquisition of desired properties such as enhanced recombination, gene
XX copy number or gene reductivity and capacity for expression and/or
XX secretion of proteins or secondary metabolites. The method is further
XX useful for predicting efficacy of a drug in treating viral infection,
XX plant genome shuffling, micropore manipulation, producing transgenic
XX animals, improvement of overexpressed genes for a desired phenotype and
XX reiterative pooling and breeding of higher organisms. The cells evolved
XX by the method are useful in molecular genetics. This sequence represents
XX an E. coli recA protein clone used in the method of the invention.

XX Sequence 358 AA;

Query Match 99.2%; Score 1665; DB 8; Length 358;

Best Local Similarity 99.1%; Pred. No. 5.9e-154;

Matches 332; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AIDENKQALAAALGQIEKQFGKGSIMRLGDRSDVKTISTGSLSLDIALGAGGLPMGR 60

Db 7 AIDENKQALAAALGQIEKQFGKGSIMRLGDRSDVKTISTGSLSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDDNLLCSQPD 120

Db 67 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDDNLLCSQPD 126

Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTTPKAEIEGIGDSHMGLAARMMSQAMRKLAG 180

Db 127 TGEQALEICDALARSGAVDVIVDSVAALTTPKAEIEGIGDSHMGLAARMMSQAMRKLAG 186

Qy 181 NLKQSNLTLLIFINQIRMKIGWFGNPNPTTTCGNALKFYASVRLDIRRIGAVKEGENVYGS 240

Db 187 NLKQSNLTLLIFINQIRMKIGWFGNPNPTTTCGNALKFYASVRLDIRRIGAVKEGENVYGS 246

Qy 241 ETRVKVKNKIAAPFKQAEFQILYEGEINFYGVLDLGVKEKLIBKAGAWYSYKEKIGQ 300

Db 247 ETRVKVKNKIAAPFKQAEFQILYEGEINFYGVLDLGVKEKLIBKAGAWYSYKEKIGQ 306

QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 335
DB 307 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 341

RESULT 15

AAW64215
ID AAW64215 standard; protein; 358 AA.

XX AC AAW64215;

XX XX 28-APR-1999 (first entry)

XX DE Hyperrecombinogenic variant recA protein clone 4.

XX XX Escherichia coli; recA protein; hyperrecombinogenic variant; Minshall;
KW recursive sequence recombination; evolution.

XX OS Escherichia coli.
XX OS Synthetic.

XX PN WO9831837-A1.

XX PD 23-JUL-1998.

XX XX 16-JAN-1998; 98WO-US000852.

XX PF 17-JAN-1997; 97US-0035054P.

XX PR (MAXY-) MAXYGEN INC.

XX PA Delcardayre SB, Tobin MB, Stemmer WPC, Ness JE, Minshall J;
PI Patten P;

XX DR WPI; 1998-427565/60.
XX DR N-ESDB; AAW44287.

XX PT Evolution of whole cells and organisms by recursive DNA sequence
PT recombination in cells to evolve cells having acquired desired function,
PT useful in methods for predicting the efficacy of a drug in treating viral
PT or pathogenic infections.

XX PS Example 1; Fig 13; 125pp; English.

XX CC The present invention provides methods employing iterative cycles of
CC recombination and selection/screening for evolution of whole cells and
CC organisms toward acquisition of desired properties, e.g. enhanced
CC recombinogenicity, genome copy number, and capacity for expression and/or
CC secretion of proteins and secondary metabolites. The present sequence
CC represents a hyperrecombinogenic variant of a recA protein (see
CC AAW64213), from an example of the present invention

XX SQ Sequence 358 AA;

Query Match 99.2%; Score 1664; DB 2; Length 358;
Best Local Similarity 98.8%; Pred. No. 7.4e-154;
Matches 331; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRDMVKTISTGSLSLDIALGAGGLPMGR 60

DB 7 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRDMVKTISTGSLSLDIALGAGGLPMGR 66

QY 61 IVEIYGPESSKTTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120

DB 67 IVEIYGPESSKTTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126

QY 121 TGEQALEICDALARSGAVDVIVVDSVAALTTPKABIEGEGISHMGLAARMMSQAMRKLKAG 180

DB 127 TGEQALEICDALARSGAVDVIVVDSVAALTTPKABIEGEGISHMGLAARMMSQAMRKLKAG 186

QY 181 NLKQSNLTLLIPINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240

DB 187 NLKQSNLTLLIPINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 246

QY 241 ETRVKKVVKNKIAAPFKQAEFQIILYGEGINFYGELVDLGVKEKLIKKAGAWSYKGEKIGQ 300
DB 247 ETRVKKVVKNKIAAPFKQAEFQIILYGEGINFYGELVDLGVKEKLIKKAGAWSYKGEKVGQ 306

QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 335
DB 307 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 341

Search completed: February 16, 2006, 01:18:52
Job time : 133 secs

This Page Blank (uspto)

R.Morimatsu, K.; Horii, T.
Eur. J. Biochem. 228, 772-778, 1995
A;Title: The DNA-binding site of the RecA protein. Photochemical cross-linking of Tyr103
A;Reference number: S69129; MUID:95255284; PMID:7737176
A;Accession: S69129
A;Status: preliminary
A;Molecule type: protein
A;Residues: 90-108,180-184 <MOR2>
A;Cross-references: UNIPARC:UPI00001748AE
R;Gardner, R.V.; Voloshin, O.N.; Camerini-Otero, R.D.
Eur. J. Biochem. 233, 419-425, 1995
A;Title: The identification of the single-stranded DNA-binding domain of the Escherichia
A;Reference number: S63979; MUID:96067680; PMID:7588783
A;Accession: S63979
A;Molecule type: protein
A;Residues: 'XX',187-190,192-194 <GAR>
A;Cross-references: UNIPARC:UPI00001748AF
R;Yu, X.; Egelman, E.H.
submitted to the Brookhaven Protein Data Bank, December 1996
A;Reference number: A67277; PDB:2REC
A;Contents: annotation; X-ray crystallography, 2.3 angstroms, residues 4-157;166-195;211
R;Aihara, H.; Ito, Y.; Kurumizaka, H.; Terada, T.; Yokoyama, S.; Shibata, T.
submitted to the Brookhaven Protein Data Bank, January 1997
A;Reference number: A67455; PDB:1AA3
A;Contents: annotation; conformation by (1)H- and (15)N-NMR, residues 269-331
C;Genetics:
A;Gene: recA
A;Map position: 58 min
C;Function:
A;Description: plays an essential role in homologous recombination, in induction of the
C;Superfamily: recombination protein recA
C;Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop;
F;67-74/Region: nucleotide-binding motif A (P-loop)
F;141-146/Region: nucleotide-binding motif B
F;73/Binding site: ATP (Lys) #status predicted

Query Match 99.8%; Score 1674; DB 1; Length 353;
Best Local Similarity 99.7%; Pred. No. 3.6e-106;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGOIEKQFGKGSIMRLGEDRSMVDVTISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGOIEKQFGKGSIMRLGEDRSMVDVTISTGSLSLDIALGAGGLPMGR 61

Qy 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSGKTTLTQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 121

Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGLAARMMSQAWRKLAG 180
Db 122 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGLAARMMSQAWRKLAG 181

Qy 181 NLKQNTLLIFINQIRMKIGWFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 182 NLKQNTLLIFINQIRMKIGWFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 241

Qy 241 ETRVKVKNKIAAPFKQAEFQILYGEINFGELVDLGKVKLEIKAGAWTSYKGEKIGQ 300
Db 242 ETRVKVKNKIAAPFKQAEFQILYGEINFGELVDLGKVKLEIKAGAWTSYKGEKIGQ 301

Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSNP 335
Db 302 GKANATAWLKONPETAKEIEKKVRELLLSNPNSNP 336

RESULT 3
H85917
hypochemical protein recA [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: H85917
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouais, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: H85917
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-353 <STO>
A;Cross-references: UNIPARC:UPI00001658E3; GB:A8005174; NID:gl2517139; PIDN:AAGS7804.1;
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: recA
C;Superfamily: recombination protein recA

Query Match 99.0%; Score 1662; DB 2; Length 353;
Best Local Similarity 99.1%; Pred. No. 2.3e-105;
Matches 332; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGOIEKQFGKGSIMRLGEDRSMVDVTISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGOIEKQFGKGSIMRLGEDRSMVDVTISTGSLSLDIALGAGGLPMGR 61

Qy 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120

DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: D91073
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-353 <HAY>
A;Cross-references: UNIPROT:P03017; UNIPARC:UPI0000112C1C; GB:BA000007; PIDN:BA036979.1;
A;Experimental source: strain O157:H7, substrain RMD 0509952
C;Genetics:
A;Gene: EC63556
C;Superfamily: recombination protein recA

Query Match 99.8%; Score 1674; DB 2; Length 353;
Best Local Similarity 99.7%; Pred. No. 3.6e-106;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGOIEKQFGKGSIMRLGEDRSMVDVTISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGOIEKQFGKGSIMRLGEDRSMVDVTISTGSLSLDIALGAGGLPMGR 61

Qy 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSGKTTLTQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 121

Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGLAARMMSQAWRKLAG 180
Db 122 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGLAARMMSQAWRKLAG 181

Qy 181 NLKQNTLLIFINQIRMKIGWFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 182 NLKQNTLLIFINQIRMKIGWFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 241

Qy 241 ETRVKVKNKIAAPFKQAEFQILYGEINFGELVDLGKVKLEIKAGAWTSYKGEKIGQ 300
Db 242 ETRVKVKNKIAAPFKQAEFQILYGEINFGELVDLGKVKLEIKAGAWTSYKGEKIGQ 301

Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSNP 335
Db 302 GKANATAWLKONPETAKEIEKKVRELLLSNPNSNP 336

RESULT 3
H85917
hypochemical protein recA [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: H85917
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouais, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: H85917
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-353 <STO>
A;Cross-references: UNIPARC:UPI00001658E3; GB:A8005174; NID:gl2517139; PIDN:AAGS7804.1;
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: recA
C;Superfamily: recombination protein recA

Query Match 99.0%; Score 1662; DB 2; Length 353;
Best Local Similarity 99.1%; Pred. No. 2.3e-105;
Matches 332; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGOIEKQFGKGSIMRLGEDRSMVDVTISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGOIEKQFGKGSIMRLGEDRSMVDVTISTGSLSLDIALGAGGLPMGR 61

Qy 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120

Db 62 IVEIYGPESSGKTTLTLOVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 121

Qy 121 TGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGIDGSHMGLAARMMWSQAMRKLKAG 180

Db 122 TGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGIDGSHMGLAARMMWSQAMRKLKAG 181

Qy 181 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240

Db 182 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 241

Qy 241 ETRVKVVKNTAAAPFKQAEFQILYEGINFGVGLVDLGVEKLEKAGAWSYKGEKIGQ 300

Db 242 ETRVKVVKNTAAAPFKQAEFQILYEGINFGVGLVDLGVEKLEKAGAWSYKGEKIGQ 301

Qy 301 GKANATAWLKNDPETAKEIEKKVRELLLSNPNSTP 335

Db 302 GKANATAWLKNDPETAKEIEKKVRELLLSNPNSTP 336

RESULT 4

AH0843

ReCA protein [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain CT18

C/Species: Salmonella enterica subsp. enterica serovar Typhi

A/Note: this species has also been called Salmonella typhi

C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C/Accession: AH0843

R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moulé, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.

A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A/Reference number: AB0502; MUID:2153494; PMID:11677608

A/Accession: AH0843

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-353 <PAR>

A/Cross-references: UNIPARC:UPI000005A334; GB:AL513382; PIDN:CAD05935.1; PTD:g16503906;

C/Genetics:

A:Gene: STY2950

C/Superfamily: recombination protein recA

Query Match 97.6%; Score 1637; DB 2; Length 353;

Best Local Similarity 97.0%; Pred. No. 1.2e-103;

Matches 325; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLSDIALGAGGLPMGR 60

Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLSDIALGAGGLPMGR 61

Qy 61 IVEIYGPESSGKTTLTLOVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120

Db 62 IVEIYGPESSGKTTLTLOVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 121

Qy 121 TGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGIDGSHMGLAARMMWSQAMRKLKAG 180

Db 122 TGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGIDGSHMGLAARMMWSQAMRKLKAG 181

Qy 181 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240

Db 182 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 241

Qy 241 ETRVKVVKNTAAAPFKQAEFQILYEGINFGVGLVDLGVEKLEKAGAWSYKGEKIGQ 300

Db 242 ETRVKVVKNTAAAPFKQAEFQILYEGINFGVGLVDLGVEKLEKAGAWSYKGEKIGQ 301

Qy 301 GKANATAWLKNDPETAKEIEKKVRELLLSNPNSTP 335

Db 302 GKANATAWLKNDPETAKEIEKKVRELLLSNPNSTP 336

RESULT 5

AG0401

RecA protein [imported] - Yersinia pestis (strain CO92)

C/Species: Yersinia pestis

C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C/Accession: AG0401

R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tithball, R.W.; Holden, M.T.G.; Prentice, M.F.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,

Nature 413, 523-527, 2001

A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A/Reference number: AB0001; MUID:21470413; PMID:11586360

A/Accession: AG0401

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-356 <KUR>

A/Cross-references: UNIPROT:P37850; UNIPARC:UPI0000165AF8; GB:AL590842; PIDN:CAC92539.1

C/Genetics:

A:Gene: recA

C/Superfamily: recombination protein recA

Query Match 91.2%; Score 1530; DB 2; Length 356;

Best Local Similarity 91.6%; Pred. No. 2.1e-96;

Matches 304; Conservative 15; Mismatches 13; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLSDIALGAGGLPMGR 60

Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLSDIALGAGGLPMGR 61

Qy 61 IVEIYGPESSGKTTLTLOVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120

Db 62 IVEIYGPESSGKTTLTLOVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 121

Qy 121 TGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGIDGSHMGLAARMMWSQAMRKLKAG 180

Db 122 TGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGIDGSHMGLAARMMWSQAMRKLKAG 181

Qy 181 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240

Db 182 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 241

Qy 241 ETRVKVVKNTAAAPFKQAEFQILYEGINFGVGLVDLGVEKLEKAGAWSYKGEKIGQ 300

Db 242 ETRVKVVKNTAAAPFKQAEFQILYEGINFGVGLVDLGVEKLEKAGAWSYKGEKIGQ 301

Qy 301 GKANATAWLKNDPETAKEIEKKVRELLLSNPN 332

Db 302 GKANATYLNKENPAIAAELDKLREMLNGN 333

RESULT 6

S31481

recombination protein recA - Enterobacter agglomerans

N/Alternate names: recombinase A

C/Species: Enterobacter agglomerans

C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004

C/Accession: S31481

R/Rappold, C.S.J.; Klingmueller, W.

submitted to the EMBL Data Library, January 1993

A/Description: Cloning and sequencing of the recA gene from Enterobacter agglomerans 33

A/Reference number: S31480

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-354 <RAP>

A/Cross-references: UNIPROT:P33037; UNIPARC:UPI000016EC3A; GB:L03291; EMBL:Z19517; NID:

C/Genetics:

A:Gene: recA

C/Superfamily: recombination protein recA

C/Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop

F/67-74/Region: nucleotide-binding motif A (P-loop)

F/141-146/Region: nucleotide-binding motif B

F/73/Binding site: ATP (Lys) #status predicted

Query Match 91.1%; Score 1528; DB 2; Length 354;

Best Local Similarity 91.8%; Pred. No. 2.8e-96;
Matches 302; Conservative 16; Mismatches 11; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDVKTISTGSLSDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDVKTISTGSLSDIALGAGGLPMGR 61

Qy 61 IVEIYGPESSGKTTTLQVIAAAREGKTCFIDAEHALDPIYAKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPESSGKTTTLQVIAAAREGKTCFIDAEHALDPIYAKLGVDIDNLLCSQPD 121

Qy 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIGSHMGLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIGSHMGLAARMMSQAMRKLKAG 181

Qy 181 NLKQSNLTLLIFINQIRMKIGWFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 182 NLKQSNLTLLIFINQIRMKIGWFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 241

Qy 241 ETRVKVVKNTAAAPFKQAEFOILYCEGINFYGELVDLGVKELIEKAGAWYSYKGEKIGQ 300
Db 242 ETRVKVVKNTAAAPFKQAEFOILYCEGINFYGELVDLGVKELIEKAGAWYSYKGEKIGQ 301

Qy 301 GKANATAWLKNPETAKEIEKKVRELLLS 329
Db 302 GKANSNLYKENPKVAELDKKLRDMLLS 330

RESULT 7
S37586
recombination protein reca - Yersinia pestis
N:Alternate names: recombinase A
C:Species: Yersinia pestis
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
R:Kryukov, V.M.; Suchkov, I.Y.; Sazykin, I.S.; Mishankin, B.N.
submitted to the EMBL Data Library, October 1993
A:Description: Complete nucleotide sequence of Yersinia pestis reca gene.
A:Reference number: S37586
A:Accession: S37586
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-356 <KRY>
A:Cross-references: UNIPROT:P37858; UNIPARC:UPI0000170617; EMBL:X75336; NID:g406793; PID:37586

Query Match 90.9%; Score 1526; DB 2; Length 356;
Best Local Similarity 91.6%; Pred. No. 3.9e-96;
Matches 304; Conservative 15; Mismatches 13; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDVKTISTGSLSDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDVKTISTGSLSDIALGAGGLPMGR 61

Qy 61 IVEIYGPESSGKTTTLQVIAAAREGKTCFIDAEHALDPIYAKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPESSGKTTTLQVIAAAREGKTCFIDAEHALDPIYAKLGVDIDNLLCSQPD 121

Qy 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIGSHMGLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIGSHMGLAARMMSQAMRKLKAG 181

Qy 181 NLKQSNLTLLIFINQIRMKIGWFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 182 NLKQSNLTLLIFINQIRMKIGWFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 241

Qy 241 ETRVKVVKNTAAAPFKQAEFOILYCEGINFYGELVDLGVKELIEKAGAWYSYKGEKIGQ 300

Db 242 ETRVKVVKNTAAAPFKQAEFOILYCEGINFYGELVDLGVKELIEKAGAWYSYKGEKIGQ 301

Qy 301 GKANATAWLKNPETAKEIEKKVRELLLSNPN 332
Db 302 GKANSNLYKENPANAELDKKLRMLNGN 333

RESULT 8
R0EBPM
recombination protein reca - Proteus mirabilis
N:Alternate names: recombinase A
C:Species: Proteus mirabilis
C:Date: 30-Sep-1990 #sequence_revision 30-Sep-1990 #text_change 09-Jul-2004
C:Accession: S04606
R:Akaboshi, E.; Yip, M.L.R.; Howard-Flanders, P.
Nucleic Acids Res. 17, 4390, 1989
A:Title: Nucleotide sequence of the reca gene of Proteus mirabilis.
A:Reference number: S04606; MUID:89296502; PMID:2544862
A:Accession: S04606
A:Molecule type: DNA
A:Residues: 1-355 <AKA>
A:Cross-references: UNIPROT:P11406; UNIPARC:UPI000016FD85; GB:X14870; NID:g45631; PIDN:C37586

Query Match 89.4%; Score 1500; DB 1; Length 355;
Best Local Similarity 89.1%; Pred. No. 2.2e-94;
Matches 294; Conservative 22; Mismatches 14; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDVKTISTGSLSDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDVKTISTGSLSDIALGAGGLPMGR 61

Qy 61 IVEIYGPESSGKTTTLQVIAAAREGKTCFIDAEHALDPIYAKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPESSGKTTTLQVIAAAREGKTCFIDAEHALDPIYAKLGVDIDNLLCSQPD 121

Qy 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIGSHMGLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIGSHVGLAARMMSQAMRKLKAG 181

Qy 181 NLKQSNLTLLIFINQIRMKIGWFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 182 NLKQSNLTLLIFINQIRMKIGWFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 241

Qy 241 ETRVKVVKNTAAAPFKQAEFOILYCEGINFYGELVDLGVKELIEKAGAWYSYKGEKIGQ 300
Db 242 ETRVKVVKNTAAAPFKQAEFOILYCEGINFYGELVDLGVKELIEKAGAWYSYKGEKIGQ 301

Qy 301 GKANATAWLKNPETAKEIEKKVRELLSN 330
Db 302 GKANSNLYKEHPMYNLTNKLREMLNH 331

RESULT 9
R82310
reca protein VC0543 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: B82310
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, P. I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: E82310
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-412 <HEI>
A;Cross-references: UNIPARC:UPI0000164B55; GB:AB004140; GB:AE003852; NID:g9654965; PIDN:
A;Experimental source: serogroup O1; strain N16961; biotype E1 Tor
C;Genetics:
A;Gene: VC0543
A;Map position: 1
C;Superfamily: recombination protein recA

Query Match 85.3%; Score 1432; DB 2; Length 412;
Best Local Similarity 84.0%; Pred. No. 1.1e-89;
Matches 278; Conservative 32; Mismatches 21; Indels 0; Gaps 0;

QY 2 IDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDVKTISTGSLSLDIALGAGGLPMGRI 61
DB 59 MDENKQKALAAALGOIEKQFGKSGIMRLGDNRAVDVEITSGSLSLDIALGAGGLPMGRI 118
QY 62 VEIYGPSSGKTTTLTQVIAAAQREGKTCAPIDAHAHALDPIYARKLGVDIDNLLCSQPD 121
DB 119 VEIYGPSSGKTTTLTLELIAAAQREGKTCAPIDAHAHALDPIYARKLGVDIDNLLVSQPD 178
QY 122 GEOALEICDALARSGAVDIVVDSVAALTPKAEITEGISHMGHGLAARMMSQAMRKLAGN 181
DB 179 GEOALEICDALARSGAVDIVVDSVAALTPKAEITEGISHMGHGLAARMMSQAMRKLTGN 238
QY 182 LKQSNLTLLIFINQIRMKIGWFGNPGPTTGGNALKFYASVRLDIRRIGAVKEGVVNSE 241
DB 239 LKQSNLCWCFINQIRMKIGWFGNPGPTTGGNALKFYASVRLDIRRIGAVKEGVVNSE 298
QY 242 TRVVKVKNKIAAPFKQAEFQILYEGEINFYGVGELVDLGVEKLEIKAGAWSYKGEKIGQG 301
DB 299 TRIKVVKNKIAAPFKQAEFQILYEGEINFYGVGELVDLGVEKLEIKAGAWSYKGEKIGQG 358
QY 302 KANATAWLKDNPTAKIEKKVRELLLSNP 332
DB 359 KANAYKYLENPEIAKTLDDKKREMLLNPN 389

RESULT 10
S46274
recombination protein recA - Vibrio cholerae
N;Alternate names: recombinase A
C;Species: Vibrio cholerae
C;Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 02-Feb-2001
C;Accession: S46274
R;Scroeder, U.H.; Lech, A.J.; Manning, P.A.
Mol. Gen. Genet. 244, 295-302, 1994
A;Title: Gene sequence of recA(+) and construction of recA mutants of Vibrio cholerae.
A;Reference number: S46274; MUID:94335880; PMID:8058040
A;Accession: S46274
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-354 <STR>
A;Cross-references: UNIPARC:UPI0000170586; EMBL:X71969; PIDN:g530269; NID:g530269; PIDN:
C;Genetics:
A;Gene: recA
C;Superfamily: recombination protein recA
C;Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop;
F;65-72/Region: nucleotide-binding motif A (P-loop)
F;139-144/Region: nucleotide-binding motif B
F;71/Binding site: ATP (Lys) #status predicted

Query Match 84.2%; Score 1413; DB 2; Length 354;
Best Local Similarity 82.8%; Pred. No. 1.7e-88;
Matches 274; Conservative 32; Mismatches 25; Indels 0; Gaps 0;

QY 2 IDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDVKTISTGSLSLDIALGAGGLPMGRI 61
DB 1 MDENKQKALAAALGOIEKQFGKSGIMHLDGNRAVDVEITSGSLSLDITLGGGLPMGRI 60

QY 62 VEIYGPSSGKTTTLTQVIAAAQREGKTCAPIDAHAHALDPIYARKLGVDIDNLLCSQPD 121
DB 61 VEIYGPSSGKTTTLTLELIAAAQREGKTCAPIDAHAHALDPIYARKLGVDIDNLLVSQPD 120
QY 122 GEOALEICDALARSGAVDIVVDSVAALTPKAEITEGISHMGHGLAARMMSQAMRKLAGN 181
DB 121 GEOALEICDALARSGAVDIVVDSVAALTPKAEITEGISHMGHGLAARMMSQAMRKLTGN 180
QY 182 LKQSNLTLLIFINQIRMKIGWFGNPGPTTGGNALKFYASVRLDIRRIGAVKEGVVNSE 241
DB 181 LKQSNLCWCFINQIRMKIGWFGNPGPTTGGNALKFYASVRLDIRRIGAVKEGVVNSE 240
QY 242 TRVVKVKNKIAAPFKQAEFQILYEGEINFYGVGELVDLGVEKLEIKAGAWSYKGEKIGQG 301
DB 241 TRIKVVKNKIAAPFKQAEFQILYEGEINFYGVGELVDLGVEKLEIKAGAWSYKGEKIGQG 300
QY 302 KANATAWLKDNPTAKIEKKVRELLLSNP 332
DB 301 KANACKYLENPEIAKTLDDKKREMLLNPN 331

RESULT 11
QJ1461
recombination protein recA - Vibrio anguillarum
N;Alternate names: recombinase A
C;Species: Vibrio anguillarum
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
C;Accession: QJ1461
R;Tolmasek, M.E.; Gammie, A.E.; Cross, J.H.
Gene 110, 41-48, 1992
A;Title: Characterization of the recA gene of Vibrio anguillarum.
A;Reference number: QJ1461; MUID:92184113; PMID:1544576
A;Accession: QJ1461
A;Molecule type: DNA
A;Residues: 1-348 <TOL>
A;Cross-references: UNIPROT:P26348; UNIPARC:UPI0000133544; GB:M80525; NID:g155256; PIDN:
A;Experimental source: strains 775 and 531A
C;Genetics:
A;Gene: recA
C;Function:
A;Description: plays an essential role in homologous recombination, in induction of the
C;Superfamily: recombination protein recA
C;Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop;
F;65-72/Region: nucleotide-binding motif A (P-loop)
F;139-144/Region: nucleotide-binding motif B
F;256-279/Region: ATP binding #status predicted
F;71/Binding site: ATP (Lys) #status predicted

Query Match 83.9%; Score 1408.5; DB 2; Length 348;
Best Local Similarity 83.3%; Pred. No. 3.4e-88;
Matches 275; Conservative 31; Mismatches 23; Indels 1; Gaps 1;

QY 2 IDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDVKTISTGSLSLDIALGAGGLPMGRI 61
DB 1 MDENKQKALAAALGOIEKQFGKSGIMRLGDNRTMDVETISTGSLSLDIALGAGGLPMGRI 60
QY 62 VEIYGPSSGKTTTLTQVIAAAQREGKTCAPIDAHAHALDPIYARKLGVDIDNLLCSQPD 121
DB 61 VEIYGPSSGKTTTLTLELIAAAQREGKTCAPIDAHAHALDPIYARKLGVDIDNLLVSQPD 120
QY 122 GEOALEICDALARSGAVDIVVDSVAALTPKAEITEGISHMGHGLAARMMSQAMRKLAGN 181
DB 121 GEOALEICDALARSGAIDVIDVDSVAALTPKAEITEGISHMGHGLAARMMSQAMRKLTGN 180
QY 182 LKQSNLTLLIFINQIRMKIGWFGNPGPTTGGNALKFYASVRLDIRRIGAVKEGVVNSE 241
DB 181 LKQSNLCWCFINQIRMKIGWFGNPGPTTGGNALKFYASVRLDIRRIGAVKEGVVNSE 240
QY 242 TRVVKVKNKIAAPFKQAEFQILYEGEINFYGVGELVDLGVEKLEIKAGAWSYKGEKIGQG 301
DB 241 TRIKVVKNKIAAPFKQADTQILYGGQFNRREGELVDLGVEKLEIKAGAWSYKGEKIGQG 300
QY 302 KANATAWLKDNPTAKIEKKVRELLLSNP 331

||||| :||:| | :|:|:| | |
Db 301 KANACKFLRENPAPAAWALDTKLEMLL-NP 329

RESULT 12
A49929
recombination protein recA - Haemophilus influenzae (strain Rd KW20)
N:Alternate names: rec-1 protein; recombinase A
C:Species: Haemophilus influenzae
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 09-Jul-2004
C:Accession: A49929; A64080
R:Zultzy, J.J.; Barcak, G.J.
J: Bacteriol. 175, 7269-7281, 1993
A:Title: Structural organization, nucleotide sequence, and regulation of the Haemophilus
A:Reference number: A49929; MUID:94042901; PMID:8228674
A:Accession: A49929
A:Molecule type: DNA
A:Residues: 1-354 <ZUL>
A:Cross-references: UNIPROT:P43705; UNIPARC:UPI00001334E8; GB:L07529; NID:g305381; PIDN:
A:Experimental source: Rd strain KW20
R:Flaeschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64080; MUID:95350630; PMID:7542800
A:Accession: A64080
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-354 <TIGR>
A:Cross-references: UNIPARC:UPI00001334E8; GB:U32741; GB:L42023; NID:g1573582; PIDN:NAC2
A:Experimental source: Rd strain KW20
C:Genetics:
A:Function:
A:Description: plays an essential role in homologous recombination, in induction of the
C:Superfamily: recombination protein recA
C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop;
F:167-74/Region: nucleotide-binding motif A (P-loop)
F:141-146/Region: nucleotide-binding motif B
F:73/Binding site: ATP (lys) #status predicted

Query Match 79.1%; Score 1328; DB 2; Length 354;
Best Local Similarity 76.1%; Pred. No. 1e-82;
Matches 252; Conservative 43; Mismatches 36; Indels 0; Gaps 0;
Qy 1 AIDENKQKALAAALQIQIEKQFGKSGIMRLGEDRSMVKTISTGSLSLDIALGAGLPMGR 60
Db 2 ATQBEKQKALAAALQIQIEKQFGKSGIMKLGDTKTLDVESISTGSLGLDVALGIGGLPMGR 61
Qy 61 IVEIYGPSSGKTTTLTVIAAQAQREGTKAFIDAEHALDPIYARKLGVDINDLLCSQPD 120
Db 62 IVEIFGPSSGKTTTLTVIAQAQAGTKAFIDAEHALDPIYAAKLGVDVKELFVSQPD 121
Qy 121 TGEQALEICDALARGAVDVIIVDSVAALTTPKAEITEGIDSHMGSLAARMWSQAMRKLAG 180
Db 122 NGEQALEICDALVRSGAIDVIIVDSVAALTTPKAEITEGDMGDSHMGSLQARLMSQALRKLTG 181
Qy 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 182 QIKRANCLVFINQIRMKIGVMFGNPETTTGGNALKFYSSVRLDIRRIGAVKEGENVVGS 241
Qy 241 ETRVKVKNKTAAPKQAEFQILYEGEINFGYELVDLGVEKLEIKAGAWYSYKGEKITGQ 300
Db 242 ETRVKVKNKLAAPPRQVDQILYEGEISKAGELLELVKHKLVKESCAWYSYNGEKIGQ 301
Qy 301 KANATAWLKONPETAKEIEKKVRELLLSNP 331
Db 302 GKANSMKRWLNENIEKSDSEARLARLAVANP 332

RESULT 13

JC5198
recombination protein recA - Aeromonas salmonicida
N:Alternate names: recombinase A
C:Species: Aeromonas salmonicida
C:Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C:Accession: JC5198
R:Umelo, E.; Noonan, B.; Trust, T.J.
Gene 175, 133-136, 1996
A:Title: Cloning, characterization and expression of the recA gene of Aeromonas salmonic
A:Reference number: JC5198; MUID:97074662; PMID:8917089
A:Accession: JC5198
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <UME>
A:Cross-references: UNIPROT:P94190; UNIPARC:UPI00001334C5; GB:U83688; NID:g1785949; PIDN:
C:Genetics:
A:Gene: recA
C:Function:
A:Description: plays an essential role in homologous recombination, in induction of the
C:Superfamily: recombination protein recA
C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop;
F:65-72/Region: nucleotide-binding motif A (P-loop)
F:139-144/Region: nucleotide-binding motif B
F:71/Binding site: ATP (lys) #status predicted

Query Match 78.6%; Score 1319; DB 2; Length 353;
Best Local Similarity 79.0%; Pred. No. 4e-82;
Matches 259; Conservative 31; Mismatches 38; Indels 0; Gaps 0;
Qy 2 IDENKQKALAAALQIQIEKQFGKSGIMRLGEDRSMVKTISTGSLSLDIALGAGLPMGR 61
Db 1 MDQNKQKALAAALQIQIEKQFGKSGIMLLGDSKTWDIEAISTGSLDVALGIGGLPCGRI 60
Qy 62 VEIYGPSSGKTTTLTVIAAQAQREGTKAFIDAEHALDPIYARKLGVDINDLLCSQPD 121
Db 61 VEIYGPSSGKTTTLTVIAEAAQKGVKCAFIADAEHALDPIYAAKLGVDVLLISQSDT 120
Qy 122 GEQALEICDALARGAVDVIIVDSVAALTTPKAEITEGIDSHMGSLAARMWSQAMRKLAGN 181
Db 121 GEQALEICDMLVRSNAVDVIIVDSVAALTTPKAEITEGEGDSHVGLQARLMSQALRKLTAN 180
Qy 182 LKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 241
Db 181 IKVANCLCIFINQIRMKIGVMFGSPETTTGGNALKFYASVRLDIRRIGAIKEGDEVVGN 240
Qy 242 TRVKVKNKTAAPKQAEFQILYEGEINFGYELVDLGVEKLEIKAGAWYSYKGEKITGQ 301
Db 241 TRVKVKNKVAAPPKQAEFQIFYGVGISKEGELVDLGVEKHLIDKAGAWYSYNGEKIQG 300
Qy 302 KANATAWLKONPETAKEIEKKVRELLLS 329
Db 301 KANVMKLFTEKNVAAEVEARLRELLLS 328

RESULT 14
RQPSNA
RecA protein PA3617 [imported] - Pseudomonas aeruginosa (strain PA01)
N:Alternate names: recombinase A
C:Species: Pseudomonas aeruginosa
C:Date: 30-Sep-1990 #sequence_revision 30-Sep-1990 #text_change 09-Jul-2004
C:Accession: S06265; S10458; A49854; JN0304; G83192
R:Sano, Y.; Kageyama, M.
Mol. Gen. Genet. 208, 412-419, 1987
A:Title: The sequence and function of the recA gene and its protein in Pseudomonas aerug
A:Reference number: S06265; MUID:88038334; PMID:2823059
A:Accession: S06265
A:Molecule type: DNA
A:Residues: 1-346 <SAN1>
A:Cross-references: UNIPROT:P08280; UNIPARC:UPI0000133516; GB:X05691; NID:g45413; PIDN:C
A:Note: part of this sequence, including the amino end of the mature protein, was confil
R:Zaitsev, E.N.; Krjukov, V.M.; Kuzmin, N.P.; Alekseev, A.A.; Lanzov, V.A.
submitted to the EMBL Data Library, March 1990
A:Description: Restriction polymorphism and nucleotide sequence substitutions in the rec

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM protein - protein search, using sw model
Run on: February 16, 2006, 01:14:35 ; Search time 152 Seconds
(without alignments)
1554.947 Million cell updates/sec
Title: US-10-733-782-3
Perfect score: 1678
Sequence: 1 AIDENKOKALAAALGQIEKQ.....AKEIEKKVRELLSNPNSTP 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1674	99.8	352	1 RECA_ECO57	P0a7g8 escherichia
2	1674	99.8	352	1 RECA_ECOL6	P0a7g7 escherichia
3	1674	99.8	352	1 RECA_ECOLI	P0a7g6 escherichia
4	1674	99.8	352	1 RECA_SHIFL	P0a7g9 shigella fl
5	1667	99.3	352	1 RECA_SHISO	Q8zff6 shigella so
6	1637	97.6	352	1 RECA_SALTI	P65978 salmonella
7	1637	97.6	352	1 RECA_SALTY	P65977 salmonella
8	1637	97.6	353	2 Q57KU4 SALCH	Q57ku4 salmonella
9	1635	97.4	353	2 Q5PFL5 SALPA	Q5pfl5 salmonella
10	1543	92.0	355	1 RECA_PHOLL	Q7n7a6 photorhabdu
11	1533	91.4	357	1 RECA_ERWCT	Q6dl88 erwinia car
12	1530	91.2	355	1 RECA_YERPE	P37858 yersinia pe
13	1530	91.2	356	1 RECA_YERPS	Q66e70 yersinia ps
14	1528	91.1	353	1 RECA_ENTAG	P33037 enterobacte
15	1523	90.8	353	1 RECA_SERMA	P17479 serratia ma
16	1520	90.6	358	1 RECA_XENBV	P96185 xenorhabdus
17	1519	90.5	355	1 RECA_SODGL	P62220 sodalis glo
18	1509	89.9	355	2 Q9YI31 GENTR	Q9yil3 primary end
19	1500	89.4	354	1 RECA_PROMI	P11406 proteus mir
20	1497	89.2	342	1 RECA_ERWCA	P26344 erwinia car
21	1497	89.2	357	1 RECA_SHEON	Q8eb80 shewanella
22	1496	89.2	358	1 RECA_XENNE	Q8x5p5 xenorhabdus
23	1482	88.3	325	1 RECA_PROVU	P26346 proteus vul
24	1433	85.4	354	1 RECA_VIBCH	P45383 vibrio chol
25	1425	84.9	347	1 RECA_VIBPA	Q871r1 vibrio para
26	1414	84.3	348	1 RECA_VIBNA	Q6xxz07 vibrio natr
27	1414	84.3	349	1 RECA_VIBVU	Q8dc51 vibrio vuln
28	1414	84.3	349	1 RECA_VIBVY	Q7mhr4 vibrio vuln
29	1412	84.1	348	2 Q5E7G6 VIBFI	Q5e7g6 vibrio fisc
30	1408.5	83.9	348	1 RECA_VIBAN	P26348 vibrio angu
31	1400	83.4	352	1 RECA_PHOPR	Q61mu2 photobacter

32	1396	83.2	346	2 Q9S4R6 VIBCH	Q9s4r6 vibrio chol
33	1396	83.2	346	2 Q9R2W9 VIBCH	Q9r2w9 vibrio chol
34	1349	80.4	354	1 RECA_PAGMU	P95526 pasteurella
35	1328	79.1	354	1 RECA_HAEIN	P43705 haemophilus
36	1328	79.1	354	2 Q4QW72 HAE18	Q4qmw2 haemophilus
37	1323	78.8	348	1 RECA_PSEOL	Q9t9u2 pseudomonas
38	1323	78.8	372	2 Q5QUB8 IDILO	Q5qub8 idiomarina
39	1319	78.6	353	1 RECA_AERSA	P94190 aeromonas s
40	1315	78.4	351	1 RECA_MANSB	Q65qb0 manheimia
41	1313	78.2	331	1 RECA_WIGBR	Q8d2w7 wiggleswort
42	1312	78.2	352	1 RECA_ACTAC	Q91rp9 actinobacil
43	1284	76.5	368	1 RECA_PASHA	Q9rny0 pasteurella
44	1276	76.0	347	2 Q6EV36 FSEST	Q6ev36 pseudomonas
45	1275	76.0	349	2 Q41XX9 AZOVI	Q41xx9 azotobacter

ALIGNMENTS

RESULT 1

RECA_ECO57 STANDARD; PRT; 352 AA.
AC P0A7G8: P03017; P26347; P78213;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RECA protein (Recombinase A).
GN Name=recA; OrderedLocName=z4002, ECa3556;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=12074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Rotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX MEDLINE=11156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
single-stranded DNA, the ATP-dependent uptake of single-stranded
DNA by duplex DNA, and the ATP-dependent hybridization of
homologous single-stranded DNAs. It interacts with lexA causing
its activation and leading to its autocatalytic cleavage (By
similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC EMBL; AE005174; AAC57804.1; -; Genomic_DNA.
CC EMBL; BA000007; BAB36979.1; -; Genomic_DNA.
CC PIR; D91073; D91073.

```

DR HAMAP; MF_00268; -; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PR00142; RecA.
DR ProDom; PD000229; RecA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS0162; RECA_2; 1.
DR PROSITE; PS0163; RECA_3; 1.
KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
FT INIT_MET 0 By similarity.
FT NP_BIND 66 73 ATP (By similarity).
FT CONFLICT 139 139 D -> N (in Ref. 1).
SQ SEQUENCE 352 AA; 37842 MW; 989B02378EC4A02 CRC64;

Query Match 99.8%; Score 1674; DB 1; Length 352;
Best Local Similarity 99.7%; Pred. No. 1.6e-103;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSLDIALGAGGLPMGR 60
Db 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSLDIALGAGGLPMGR 60
Qy 61 IVEIYGPSSSGKTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 61 IVEIYGPSSSGKTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
Qy 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGEIGDSHMGLAARMWSQAWRKLAG 180
Db 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGEIGDSHMGLAARMWSQAWRKLAG 180
Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
Db 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
Qy 241 ETRVKVVKNTKIAAPFKQAEFQILYEGEINFGYELVDLGKVEKLEKAGWYSYKGEKIQ 300
Db 241 ETRVKVVKNTKIAAPFKQAEFQILYEGEINFGYELVDLGKVEKLEKAGWYSYKGEKIQ 300
Qy 301 GKANATAMKONPETAKEIEKKVRELLLSNPSTP 335
Db 301 GKANATAMKONPETAKEIEKKVRELLLSNPSTP 335

RESULT 2
RECA_ECOL6 STANDARD; PRT; 352 AA.
AC POA7G7; P03017; P26347; P78213;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocuNames=b2699;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
RX MEDLINE=2238234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Raeko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded

```

```

CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AE016765; AAN81704.1; -; Genomic_DNA.
CC HAMAP; MF_00268; -; 1.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR001553; RecA.
CC Pfam; PF00154; RecA; 1.
CC PRINTS; PR00142; RecA.
CC ProDom; PD000229; RecA; 1.
CC SMART; SM00382; AAA; 1.
CC TIGRFAMs; TIGR02012; tigrfam_recA; 1.
CC PROSITE; PS00321; RECA_1; 1.
CC PROSITE; PS0162; RECA_2; 1.
CC PROSITE; PS0163; RECA_3; 1.
KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
FT INIT_MET 0 By similarity.
FT NP_BIND 66 73 ATP (By similarity).
FT CONFLICT 139 139 D -> N (in Ref. 1).
SQ SEQUENCE 352 AA; 37842 MW; 989B02378EC4A02 CRC64;

Query Match 99.8%; Score 1674; DB 1; Length 352;
Best Local Similarity 99.7%; Pred. No. 1.6e-103;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSLDIALGAGGLPMGR 60
Db 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSLDIALGAGGLPMGR 60
Qy 61 IVEIYGPSSSGKTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 61 IVEIYGPSSSGKTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
Qy 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGEIGDSHMGLAARMWSQAWRKLAG 180
Db 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGEIGDSHMGLAARMWSQAWRKLAG 180
Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
Db 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
Qy 241 ETRVKVVKNTKIAAPFKQAEFQILYEGEINFGYELVDLGKVEKLEKAGWYSYKGEKIQ 300
Db 241 ETRVKVVKNTKIAAPFKQAEFQILYEGEINFGYELVDLGKVEKLEKAGWYSYKGEKIQ 300
Qy 301 GKANATAMKONPETAKEIEKKVRELLLSNPSTP 335
Db 301 GKANATAMKONPETAKEIEKKVRELLLSNPSTP 335

RESULT 3
RECA_ECOLI STANDARD; PRT; 352 AA.
AC POA7G6; P03017; P26347; P78213;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; Synonym=lexB, rech, rnmB, tif, umuB, zab;
GN OrderedLocuNames=b2699;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.

```


D -> E (in Ref. 5). Missing (in Ref. 7).		99.8%; Score 1674; DB 1; Length 352; 99.7%; Pred. No. 1.6e-103; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
112	112	1	AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVKTISTGSLSLDIALGAGGLPMGR 60
190	190	1	AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 60
FT	CONFLICT	61	IVEIYGPSSGKTTLTQVIAAAREGKTCAPIDAEHALDPIYARKLGVDDNLLCSQPD 120
FT	TURN	61	IVEIYGPSSGKTTLTQVIAAAREGKTCAPIDAEHALDPIYARKLGVDDNLLCSQPD 120
FT	HELIX	121	TGEQALEICDARSQVDVIVDSVAALTPKAEIEGIBGSHMGLAARMMSQAMKLAG 180
FT	HELIX	121	TGEQALEICDARSQVDVIVDSVAALTPKAEIEGIBGSHMGLAARMMSQAMKLAG 180
FT	TURN	181	NLKQSNLTLLIFNQIRMKIGWFGNPETTTGGNALKFYASVRLDIRRGVKEGENVVGS 240
FT	TURN	181	NLKQSNLTLLIFNQIRMKIGWFGNPETTTGGNALKFYASVRLDIRRGVKEGENVVGS 240
FT	STRAND	241	ETRVKVNKNTAAPKQAEFQLLYGEGINFYGELVDLGKVKLEIKAGWTSYKGEKIQ 300
FT	STRAND	241	ETRVKVNKNTAAPKQAEFQLLYGEGINFYGELVDLGKVKLEIKAGWTSYKGEKIQ 300
FT	TURN	301	GKANATAWLKDNPETAKEIEKKVRELLLSNPNSTP 335
FT	TURN	301	GKANATAWLKDNPETAKEIEKKVRELLLSNPNSTP 335

RESULT 4
RECA_SHIFL STANDARD; PRT; 352 AA.
AC POA7G9; P03017; P26347; P78213;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A)
GN Name=recA; OrderedLocustNames=SF2722, S2913;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BS12 / Serotype 2a;
RX MEDLINE=91109725; PubMed=2274037;
RA Zhao X.J., McEntee K.;
RT "DNA sequence analysis of the recA genes from Proteus vulgaris,
RL Erwinia carotovora, Shigella flexneri and Escherichia coli B/r.",
RN Mol. Gen. Genet. 222:369-376(1990).
[2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=2272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
through comparison with genomes of Escherichia coli K12 and O157.",
RL Nucleic Acids Res. 30:4432-4441(2002).
[3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
flexneri serotype 2a strain 2457T.",
RL Infect. Immun. 71:2775-2786(2003).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with LexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; X55553; CAB56806.1; -; Genomic DNA.
CC EMBL; AE005674; AAN44214.1; ALT_INIT; Genomic_DNA.
CC EMBL; AE016987; AAP18040.1; -; Genomic_DNA.
CC HAMAP; MF_00268; -; 1.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR001553; RecA.
CC Pfam; PF00154; RecA; 1.
CC PRINTS; PR00142; RecA.
CC PRODOM; PD000229; RecA; 1.
CC SMART; SM00382; AAA; 1.
CC TIGRFAMs; TIGR02012; tigrfam_recA; 1.
CC PROSITE; PS00321; RECA_1; 1.
CC PROSITE; PS50162; RECA_2; 1.

```

DR PROSITE; PS50163; RECA_3; 1.
KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
DR INIT MET 0 By similarity.
FT NP BIND 66 73 ATP (By similarity).
SQ SEQUENCE 352 AA; 37842 MW; 989B02378ECA4A02 CRC64;

Query Match 99.8%; Score 1674; DB 1; Length 352;
Best Local Similarity 99.7%; Pred. No. 1.6e-103;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGSDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60
DB 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGSDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60
QY 61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
QY 121 TGEQALEICDALARSGADVIVVDSVAALTPKABIEGIEGDSHMGLAARMMSQAMRKLKAG 180
DB 121 TGEQALEICDALARSGADVIVVDSVAALTPKABIEGIEGDSHMGLAARMMSQAMRKLKAG 180
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
DB 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
QY 241 ETRVKVVKNTAAAPFKQAEFQILYGEINFGYELVDLGVEKLEKAGAWTSYKGEKIGQ 300
DB 241 ETRVKVVKNTAAAPFKQAEFQILYGEINFGYELVDLGVEKLEKAGAWTSYKGEKIGQ 300
QY 301 GKANATLAWLNDPETAKEIEKKVRELLSNPNSTP 335
DB 301 GKANATLAWLNDPETAKEIEKKVRELLSNPNSTP 335

RESULT 5
RECA SHISO STANDARD; PRT; 352 AA.
AC O92FF6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA;
OS Shigella sonnei.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=624;
[1]
RN NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=KNH104S;
RA Park Y.C., Shin H.J., Kim Y.C.;
RT "Cloning and nucleotide sequence of the recA gene from Shigella sonnei
  KNH104S."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
  single-stranded DNA, the ATP-dependent uptake of single-stranded
  DNA by duplex DNA, and the ATP-dependent hybridization of
  homologous single-stranded DNAs. It interacts with LexA causing
  its activation and leading to its autocatalytic cleavage.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use as long as its content is in no way modified and this statement is not
  removed.
CC -----
CC EMBL; AF101227; AAC72856.1; -; Genomic_DNA.
CC HSPSP; P03017; 2REB.
CC SMR; O92FF6; 3-328.

```

```

DR HAWAP; MF_00268; -; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PR00142; RecA.
DR ProDom; PD000229; RecA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS00162; RECA_2; 1.
DR PROSITE; PS00163; RECA_3; 1.
KW ATP-binding; DNA damage; DNA recombination; DNA repair; DNA-binding;
KW Nucleotide-binding; SOS response.
DR INIT MET 0 By similarity.
FT NP BIND 66 73 ATP (By similarity).
SQ SEQUENCE 352 AA; 37799 MW; BEC50231893AA40C CRC64;

Query Match 99.3%; Score 1667; DB 1; Length 352;
Best Local Similarity 99.4%; Pred. No. 4.6e-103;
Matches 333; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGSDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60
DB 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGSDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60
QY 61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
QY 121 TGEQALEICDALARSGADVIVVDSVAALTPKABIEGIEGDSHMGLAARMMSQAMRKLKAG 180
DB 121 TGEQALEICDALARSGADVIVVDSVAALTPKABIEGIEGDSHMGLAARMMSQAMRKLKAG 180
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
DB 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
QY 241 ETRVKVVKNTAAAPFKQAEFQILYGEINFGYELVDLGVEKLEKAGAWTSYKGEKIGQ 300
DB 241 ETRVKVVKNTAAAPFKQAEFQILYGEINFGYELVDLGVEKLEKAGAWTSYKGEKIGQ 300
QY 301 GKANATLAWLNDPETAKEIEKKVRELLSNPNSTP 335
DB 301 GKANATLAWLNDPETAKEIEKKVRELLSNPNSTP 335

RESULT 6
RECA SALT1 STANDARD; PRT; 352 AA.
AC P65978; Q8XET0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocusNames=STY2950, t2730;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CT18;
RC MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatina M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dow L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
  enterica serovar Typhi CT18."

```

```
RL Nature 413:848-852 (2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RY DOI=10.1128/JB.185.7.2330-2337.2003;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.:
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and Cr18."
RJ J. Bacteriol. 185:2330-2337 (2003).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AL627276; CAD05935.1; -; Genomic DNA.
DR EMBL; AB016843; AA070291.1; -; Genomic_DNA.
DR HSSP; P03017; 2REB.
DR SMR; P65978; 3-328.
DR HAWAP; MF_00268; -; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR PRINTS; PS00154; RECA.
DR PRODOM; PD000229; RECA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR02012; tigrfam_reca; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS00162; RECA_2; 1.
DR PROSITE; PS00163; RECA_3; 1.
KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
FT INIT MET 0
FT NP_BIND 66 73 ATP (By similarity).
FT SEQUENCE 352 AA; 37813 MW; F24E51CB9A31B818 CRC64;
Query Match 97.6%; Score 1637; DB 1; Length 352;
Best Local Similarity 97.0%; Pred. No. 4.6e-101;
Matches 325; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
QY 1 AIDENKQALAAALGQIEKQFGKSGIMRLGEDRSMVDVTISTGSLSLDIALGAGGLPMGR 60
Db 1 AIDENKQALAAALGQIEKQFGKSGIMRLGEDRSMVDVTISTGSLSLDIALGAGGLPMGR 60
QY 61 IVEIYGPSSGKTTTLQVIAAAQREGTKAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 61 IVEIYGPSSGKTTTLQVIAAAQREGTKAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIGDSHMGGLAARMKSOAMRKLAG 180
Db 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIGDSHMGGLAARMKSOAMRKLAG 180
QY 181 NLKQNTLLIFINQIRMKIGVFGNPGPTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
Db 181 NLKQNTLLIFINQIRMKIGVFGNPGPTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
QY 241 ETRVKVKNKIAAPKQAEFQILYEGEINFGYELVDLGKELIEKAGWYSYKGEKIGQ 300
Db 241 ETRVKVKNKIAAPKQAEFQILYEGEINFGYELVDLGKELIEKAGWYSYKGEKIGQ 300
QY 301 GKANATAWLNKONPETAKEIEKRVRELLLSNQNPSTP 335
Db 301 GKANATAWLNKONPETAKEIEKRVRELLLSNQNPSTP 335
```

```
Db 301 GKANATTWLNKENPATAKEIEKRVRELLLSNQNPATP 335
RESULT 7
RECA SALTY
ID RECA SALTY STANDARD; PRT; 352 AA.
AC P65977; O8XET0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocusNames=STM2829;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856 (2001).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AE008829; AAL21709.1; -; Genomic_DNA.
DR HSSP; P03017; 2REB.
DR SMR; P65977; 3-328.
DR HAWAP; MF_00268; -; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR PRINTS; PS00154; RECA.
DR PRODOM; PD000229; RECA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR02012; tigrfam_reca; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS00162; RECA_2; 1.
DR PROSITE; PS00163; RECA_3; 1.
KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
FT INIT MET 0
FT NP_BIND 66 73 ATP (By similarity).
FT SEQUENCE 352 AA; 37813 MW; F24E51CB9A31B818 CRC64;
Query Match 97.6%; Score 1637; DB 1; Length 352;
Best Local Similarity 97.0%; Pred. No. 4.6e-101;
Matches 325; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
QY 1 AIDENKQALAAALGQIEKQFGKSGIMRLGEDRSMVDVTISTGSLSLDIALGAGGLPMGR 60
Db 1 AIDENKQALAAALGQIEKQFGKSGIMRLGEDRSMVDVTISTGSLSLDIALGAGGLPMGR 60
QY 61 IVEIYGPSSGKTTTLQVIAAAQREGTKAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 61 IVEIYGPSSGKTTTLQVIAAAQREGTKAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIGDSHMGGLAARMKSOAMRKLAG 180
Db 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIGDSHMGGLAARMKSOAMRKLAG 180
QY 181 NLKQNTLLIFINQIRMKIGVFGNPGPTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
Db 181 NLKQNTLLIFINQIRMKIGVFGNPGPTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
QY 241 ETRVKVKNKIAAPKQAEFQILYEGEINFGYELVDLGKELIEKAGWYSYKGEKIGQ 300
Db 241 ETRVKVKNKIAAPKQAEFQILYEGEINFGYELVDLGKELIEKAGWYSYKGEKIGQ 300
QY 301 GKANATAWLNKONPETAKEIEKRVRELLLSNQNPSTP 335
Db 301 GKANATTWLNKENPATAKEIEKRVRELLLSNQNPATP 335
```

```

Db      61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAFIDAETHALDPVYARKLGVDIDNLLCSQPD 120
QY      121 TGEQALICDALARSGADVIVVDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 180
Db      121 TGEQALICDALARSGADVIVVDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 180
QY      181 NLKQSNLTLLIFINQIRMKIGVMFGNPGPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db      181 NLKQSNLTLLIFINQIRMKIGVMFGNPGPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
QY      241 ETRVKKVKNKIAAPFKQAEFOILYEGINFGYELVDLGVKEKLEKAGAWNTSYNGEKIKG 300
Db      241 ETRVKKVKNKIAAPFKQAEFOILYEGINFGYELVDLGVKEKLEKAGAWNTSYNGEKIKG 300
QY      301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db      301 GKANATTWLNKENPATAKEIEKRVRELLLSNQATP 335

```

RESULT 8

```

Q57KU4_SALCH
ID Q57KU4_SALCH PRELIMINARY; PRT; 353 AA.
AC Q57KU4
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE DNA strand exchange and recombination protein with protease and
DE nuclease activity.
GN Name=RecA; OrderedLocusNames=SC2762;
OS Salmonella choleraesuis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=591;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SC-B67;
RX PubMed=15781495; DOI=10.1093/nar/gki297;
RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
RA Wang H.-S., Lee Y.-S.;
RT "The genome sequence of Salmonella enterica serovar Choleraesuis, a
RT highly invasive and resistant zoonotic pathogen."
RL Nucleic Acids Res. 33:1690-1698(2005).
DR EMBL; AE017220; AAX66668.1; -; Genomic_DNA.
DR SMR; Q57KU4; 4-329.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003684; F:damaged DNA binding; IEA.
DR GO; GO:0008094; F:DNA-dependent ATPase activity; IEA.
DR GO; GO:0017111; F:nucleoside-triphosphatase activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PR00142; RECA.
DR SMART; SM00382; AAA; 1.
DR TIGRfams; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS00162; RECA_2; 1.
DR PROSITE; PS0163; RECA_3; 1.
DR Complete proteome.
SQ SEQUENCE 353 AA; 37944 MW; ECBBAB042F16362C CRC64;

```

Query Match 97.6%; Score 1637; DB 2; Length 353;
 Best Local Similarity 97.0%; Pred. No. 4.6e-101;
 Matches 325; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

```

QY      1 AIDENKQALAAALGQIEKQFGKGSIMRLGEDRSMVDKTIISTGSLSDIALGAGGLPMGR 60
Db      2 AIDENKQALAAALGQIEKQFGKGSIMRLGEDRSMVDKTIISTGSLSDIALGAGGLPMGR 61
QY      61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAFIDAETHALDPVYARKLGVDIDNLLCSQPD 120

```

```

Db      62 IVEIYGPESSGKTTLTQVIAAAQREGKTCAFIDAETHALDPVYARKLGVDIDNLLCSQPD 121
QY      121 TGEQALICDALARSGADVIVVDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 180
Db      122 TGEQALICDALARSGADVIVVDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 181
QY      181 NLKQSNLTLLIFINQIRMKIGVMFGNPGPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db      182 NLKQSNLTLLIFINQIRMKIGVMFGNPGPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 241
QY      241 ETRVKKVKNKIAAPFKQAEFOILYEGINFGYELVDLGVKEKLEKAGAWNTSYNGEKIKG 300
Db      242 ETRVKKVKNKIAAPFKQAEFOILYEGINFGYELVDLGVKEKLEKAGAWNTSYNGEKIKG 301
QY      301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db      302 GKANATTWLNKENPATAKEIEKRVRELLLSNQATP 336

```

RESULT 9

```

Q5PF15_SALPA
ID Q5PF15_SALPA PRELIMINARY; PRT; 353 AA.
AC Q5PF15;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE RecA protein.
GN Name=RecA; OrderedLocusNames=SPA2687;
OS Salmonella paratyphi-a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=54388;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 9150 / SARB42;
RX PubMed=15531882; DOI=10.1038/ng1470;
RA McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,
RA Porwolland S., Sabo A., Meyer R., Bieri T., Ozersky P., McLellan M.,
RA Harkins C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
RA Kohlberg S., Strong C., Du F., Carter J., Kremizki C., Layman D.,
RA Leonard S., Sun H., Fulton L., Nash W., Miner T., Minx P.,
RA Delehaunty K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
RA Spieth J., Wilson R.K.;
RT "Comparison of genome degradation in Paratyphi A and Typhi, human-
RT restricted serovars of Salmonella enterica that cause typhoid."
RL Nat. Genet. 36:1268-1274(2004).
DR EMBL; CP000026; AAV78544.1; -; Genomic_DNA.
DR SMR; Q5PF15; 4-329.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003684; F:damaged DNA binding; IEA.
DR GO; GO:0008094; F:DNA-dependent ATPase activity; IEA.
DR GO; GO:0017111; F:nucleoside-triphosphatase activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PR00142; RECA.
DR ProDom; PD000229; RecA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRfams; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS00162; RECA_2; 1.
DR PROSITE; PS0163; RECA_3; 1.
DR Complete proteome.
SQ SEQUENCE 353 AA; 37962 MW; AD15A0C48E7CF6FD CRC64;

```

Query Match 97.4%; Score 1635; DB 2; Length 353;
 Best Local Similarity 96.7%; Pred. No. 6.2e-101;
 Matches 324; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

```
Qy 1 AIDENKOKALAAALGOIEKQFGKSGIMRLGDRSDVKTISTGSLSDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGOIEKQFGKSGIMRLGDRSDVKTISTGSLSDIALGAGGLPMGR 61
Qy 61 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 121
Qy 121 TGEQALBICDALARSGAVDVIVDSVAALTPKABIEIGDSHGLAARMMSQAMRKLKAG 180
Db 122 TGEQALBICDALARSGAVDVIVDSVAALTPKABIEIGDSHGLAARMMSQAMRKLKAG 181
Qy 181 NLKOSNTLLIFINQIRMKIGVMFNGPPTTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240
Db 182 NLKOSNTLLIFINQIRMKIGVMFNGPPTTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 241
Qy 241 ETRVKKVKNKIAAPFKQAEFQILYEGEGINFGYELVDLGVKEKLEKAGWYSYKGEKIGQ 300
Db 242 ETRVKKVKNKIAAPFKQAEFQILYEGEGINFGYELVDLGVKEKLEKAGWYSYKGEKIGQ 301
Qy 301 GKANATAWLKNPETAKEIEKKVRELLLSNPNSTP 335
Db 302 GKANATWLNKENPATAKEIEKRVRELLLSNQWATP 336

RESULT 10
RECA PHOLL
ID RECA PHOLL STANDARD; PRT; 355 AA.
AC Q7N7A6;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocusNames=plu1249;
OS Photorhabdus luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
OX NCBI_TaxID=141679;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=TT01;
RX MEDLINE=22957627; PubMed=14528314; DOI=10.1038/nbt886;
RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,
RA Taouric S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-F.,
RA Dassa E., Derose R., Derzelle S., Freysinet G., Gaudriault S.,
RA Medigue C., Lanolis A., Powell K., Signier P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.
RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
RT luminescens".
RL Nat. Biotechnol. 21:1307-1313(2003).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC -----
CC EMBL; BX571863; CAB13543.1; -; Genomic_DNA.
CC SMR; Q7N7A6; 4-329.
CC PhotoList; plu1249; -.
CC HAMAP; MF_00268; -; 1.
CC InterPro; IPR001553; RecA.
CC Pfam; PF00154; RecA; 1.
CC PRINTS; PR00142; RecA.
CC ProDom; PD000229; RecA; 1.
```

```
DR TIGRFAMS; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS0162; RECA_2; 1.
DR PROSITE; PS0163; RECA_3; 1.
KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
FT NP BIND 67 74 ATP (By similarity).
SQ SEQUENCE 355 AA; 36381 MW; FD659ACED827AEF9 CRC64;

Query Match 92.0%; Score 1543; DB 1; Length 355;
Best Local Similarity 91.8%; Pred. No. 8.5e-95;
Matches 303; Conservative 16; Mismatches 11; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGOIEKQFGKSGIMRLGDRSDVKTISTGSLSDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGOIEKQFGKSGIMRLGDRSDVKTISTGSLSDIALGAGGLPMGR 61
Qy 61 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 121
Qy 121 TGEQALBICDALARSGAVDVIVDSVAALTPKABIEIGDSHGLAARMMSQAMRKLKAG 180
Db 122 TGEQALBICDALARSGAVDVIVDSVAALTPKABIEIGDSHGLAARMMSQAMRKLKAG 181
Qy 181 NLKOSNTLLIFINQIRMKIGVMFNGPPTTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240
Db 182 NLKOSNTLLIFINQIRMKIGVMFNGPPTTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 241
Qy 241 ETRVKKVKNKIAAPFKQAEFQILYEGEGINFGYELVDLGVKEKLEKAGWYSYKGEKIGQ 300
Db 242 ETRVKKVKNKIAAPFKQAEFQILYEGEGINFGYELVDLGVKEKLEKAGWYSYKGEKIGQ 301
Qy 301 GKANATAWLKNPETAKEIEKKVRELLLSN 330
Db 302 GKANATTYKHEPEVATELCKLREMLLN 331

RESULT 11
RECA ERWCT
ID RECA ERWCT STANDARD; PRT; 357 AA.
AC Q6D1S8;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocusNames=ECA3369;
OS Erwinia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SCRI 1043 / ATCC BAA-672;
RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebahia M., Pritchard L., Holden M.T.G., Hyman L.J.,
RA Holvea M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagers K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.
RT "Genome sequence of the enterobacterial phytopathogen Erwinia
RT carotovora subsp. atroseptica and characterization of virulence
RT factors".
RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
```



```

CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; BX590851; CAG76267.1; -; Genomic_DNA.
CC DR SMR; Q6D1S8; 4-329.
CC DR HAMAP; MF_00268; -; 1.
CC DR InterPro; IPR003593; AAA_ATPase.
CC DR InterPro; IPR001553; RecA.
CC DR Pfam; PF00154; RecA; 1.
CC DR PRINTS; PD00142; RECA.
CC DR ProDom; PD000229; RecA; 1.
CC DR SMART; SM00382; AAA; 1.
CC DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
CC DR PROSITE; PS00321; RECA_1; 1.
CC DR PROSITE; PS0162; RECA_2; 1.
CC DR PROSITE; PS0163; RECA_3; 1.
CC DR ATP-binding; Complete proteome; DNA damage; DNA recombination;
CC DNA repair; DNA-binding; Nucleotide-binding; SOS response.
CC KW NP BIND 67 74 ATP (By similarity).
CC FT SEQUENCE 357 AA; 38284 MW; 49F1FBB8232E0098 CRC64;
CC SQ
Query Match 91.4%; Score 1533; DB 1; Length 357;
Best Local Similarity 91.3%; Pred. No. 4e-94;
Matches 303; Conservative 15; Mismatches 14; Indels 0; Gaps 0;
QY 1 AIDENKOKAALAAALGQIEKQFGKGSIMRLGSDRSMDVKTITSTGSLSLDIALGAGGLPMGR 60
DB 2 AIDENKOKAALAAALGQIEKQFGKGSIMRLGSDRSMDVKTITSTGSLSLDIALGAGGLPMGR 61
QY 61 IVEIYGPESSGKTTLTQVIAAAREGKTCAFDAEHALDPIYAKLGVDDINLLCSQPD 120
DB 62 IVEIYGPESSGKTTLTQVIAAAREGKTCAFDAEHALDPIYAKLGVDDINLLCSQPD 121
QY 121 TGEQALBICDALARSGAVDVIVDSVAALTPKABIEGIEIGDGHMLAARWMSQAMRKLKAG 180
DB 122 TGEQALBICDALTSGAVDVIVDSVAALTPKABIEGIEIGDGHMLAARWMSQAMRKLKAG 181
QY 181 NLKQSNLTLLFIQIRKMGVFMGNPETTTGGNALKFYASVRLDIRIGAVKEGENVVGS 240
DB 182 NLKQANTLLFIQIRKMGVFMGNPETTTGGNALKFYASVRLDIRIGAVKEGENVVGS 241
QY 241 ETRVKVVKNTAAAPKQAEFQILYGEINFGVGLVDLGVKELIEKAGAMYSYKGEKIGQ 300
DB 242 ETRVKVVKNTAAAPKQAEFQILYGEINFGVGLVDLGVKELIEKAGAMYSYNGDKIGQ 301
QY 301 GKANATAWLKDNPETAKEIEKKVRELLSNPN 332
DB 302 GKANACNFKLENPTISAELOKCLREMLLHKN 333
RESULT 12
RECA_YERPE
ID RECA_YERPE STANDARD; PRT; 355 AA.
AC P37858;
DT 01-OCT-1994 (Rel. 30, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name:recA; OrderedLocName:YPO3307, y0881, YPO379;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=231;
RA Kryukov V.M., Suchkov I.Y., Sazykin I.S., Mishankin B.N.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [2]

```

```

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CO-92 / Biovar Orientalis;
RC MEDLINE=21470413; PubMed=1158636; DOI=10.1038/35097083;
RA Packhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebaihia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=KIM5 / Biovar Mediaevalis;
RC MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.P., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=91001 / Biovar Mediaevalis;
RC PubMed=15368893;
RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
RA Yang H., Wang J., Huang P., Yang R.;
RT "Complete genome sequence of Yersinia pestis strain 91001, an isolate
RT avirulent to humans.";
RL DNA Res. 11:179-197(2004).
CC -1- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the recA family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; X75336; CAA53084.1; -; Genomic DNA.
CC DR EMBL; AJ414156; CAC92539.1; -; Genomic DNA.
CC DR EMBL; AE013691; AAM84465.1; -; Genomic DNA.
CC DR EMBL; AE017128; AAS60652.1; -; Genomic DNA.
CC DR PIR; AG0401; AG0401.
CC DR PIR; S37586; S37586.
CC DR HSPP; P03017; 2REB.
CC DR SMR; P37858; 3-328.
CC DR HAMAP; MF_00268; -; 1.
CC DR InterPro; IPR003593; AAA_ATPase.
CC DR InterPro; IPR001553; RecA.
CC DR Pfam; PF00154; RecA; 1.
CC DR PRINTS; PR00142; RECA.
CC DR ProDom; PD000229; RecA; 1.
CC DR SMART; SM00382; AAA; 1.
CC DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
CC DR PROSITE; PS00321; RECA_1; 1.
CC DR PROSITE; PS0162; RECA_2; 1.
CC DR PROSITE; PS0163; RECA_3; 1.
CC DR ATP-binding; Complete proteome; DNA damage; DNA recombination;
CC DNA repair; DNA-binding; Nucleotide-binding; SOS response.
CC KW INIT MET 0 0 By similarity.
CC FT NP BIND 66 73 ATP (By similarity).
CC FT CONFLICT 281 281 H -> L (in Ref. 1).
CC FT

```

```

FT CONFLICT 294 294 N -> Y (in Ref. 1).
FT CONFLICT 315 315 I -> N (in Ref. 1).
SQ SEQUENCE 355 AA; 37755 MW; F800D2D1AD32AB81 CRC64;

Query Match 91.2%; Score 1530; DB 1; Length 355;
Best Local Similarity 91.6%; Pred. No. 6.3e-94;
Matches 304; Conservative 15; Mismatches 13; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVKTISTGSLSLDIALGAGGLPMGR 60
Db 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVKTISTGSLSLDIALGAGGLPMGR 60

Qy 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCFADAEHALDPIYAKLGVDIDNLLCSQPD 120
Db 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCFADAEHALDPIYAKLGVDIDNLLCSQPD 120

Qy 121 TGEQALEICDALRSGAVDVIIVDSVAALTPKAEIEGEGSHMGLAARMWSQAMRKLKAG 180
Db 121 TGEQALEICDALRSGAVDVIIVDSVAALTPKAEIEGEGSHMGLAARMWSQAMRKLKAG 180

Qy 181 NLKQNTLLIFINQIRMKIGVFNPGPTTGGNALKFYASVRLDIRRIGAVKEGENVVGVS 240
Db 181 NLKQNTLLIFINQIRMKIGVFNPGPTTGGNALKFYASVRLDIRRIGAVKEGENVVGVS 240

Qy 241 ETRVKVVKNTAAAPFKQAEFQILYEGEINFGELVDLGKVEKLEKAGAWSYKGEKIGQ 300
Db 241 ETRVKVVKNTAAAPFKQAEFQILYEGEINFGELVDLGKVEKLEKAGAWSYKGEKIGQ 300

Qy 301 GKANATWLKONPETAKEIEKKVRELLSNPN 332
Db 301 GKANASNYLKENPAIAAELDKLREMLNGN 332

RESULT 13
RECA_YERPS
ID RECA_YERPS STANDARD; PRT; 356 AA.
AC Q66E70;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocusNames=YPTB0823;
OS Versinia pseudotuberculosis.
OC Bacteriia; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_taxID=633;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IP32953 / Serotype I;
RX PubMed=15358858; DOI=10.1073/pnas.0404012101;
RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francois V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RT "Insights into the evolution of Versinia pestis through whole-genome
RT comparison with Versinia pseudotuberculosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with LexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

```

```

DR EMBL; BX936398; CAH20063.1; -; Genomic_DNA.
DR SMR; Q66E70; 4-329.
DR HAMAP; MF_00268; -; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PR00142; RECA.
DR ProDom; PD000229; RecA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS0162; RECA_2; 1.
DR PROSITE; PS0163; RECA_3; 1.
KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
FT NP_BIND 67 74 ATP (By similarity).
SQ SEQUENCE 356 AA; 37914 MW; DIC0D90F58C8642B CRC64;

Query Match 91.2%; Score 1530; DB 1; Length 356;
Best Local Similarity 91.6%; Pred. No. 6.3e-94;
Matches 304; Conservative 15; Mismatches 13; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVKTISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVKTISTGSLSLDIALGAGGLPMGR 61

Qy 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCFADAEHALDPIYAKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSGKTTTLQVIAAAQREGKTCFADAEHALDPIYAKLGVDIDNLLCSQPD 121

Qy 121 TGEQALEICDALRSGAVDVIIVDSVAALTPKAEIEGEGSHMGLAARMWSQAMRKLKAG 180
Db 122 TGEQALEICDALRSGAVDVIIVDSVAALTPKAEIEGEGSHMGLAARMWSQAMRKLKAG 181

Qy 181 NLKQNTLLIFINQIRMKIGVFNPGPTTGGNALKFYASVRLDIRRIGAVKEGENVVGVS 240
Db 182 NLKQNTLLIFINQIRMKIGVFNPGPTTGGNALKFYASVRLDIRRIGAVKEGENVVGVS 241

Qy 241 ETRVKVVKNTAAAPFKQAEFQILYEGEINFGELVDLGKVEKLEKAGAWSYKGEKIGQ 300
Db 242 ETRVKVVKNTAAAPFKQAEFQILYEGEINFGELVDLGKVEKLEKAGAWSYKGEKIGQ 301

Qy 301 GKANATWLKONPETAKEIEKKVRELLSNPN 332
Db 302 GKANASNYLKENPAIAAELDKLREMLNGN 333

RESULT 14
RECA_ENTAG
ID RECA_ENTAG STANDARD; PRT; 353 AA.
AC P33037;
DT 01-OCT-1993 (Rel. 27, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA;
OS Enterobacter agglomerans (Erwinia herbicola) (Pantoea agglomerans).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pantoea.
OX NCBI_taxID=549;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RL Rappold C.S.J., Klingmueller W.;
CC Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with LexA causing
CC its activation and leading to its autocatalytic cleavage.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL; L03291; AAA91766.1; -; Genomic_DNA.

DR PIR; S31481; S31481.

DR HSSP; P03017; 2REB.

DR SMR; P33037; 3-328.

DR HAMAP; MF 00268; -; 1.

DR InterPro; IPR003593; AAA ATPase.

DR InterPro; IPR001553; RecA.

DR Pfam; PF00154; RecA; 1.

DR PRINTS; PR00142; RECA.

DR ProDom; PD000229; RecA; 1.

DR SMART; SM00382; AAA; 1.

DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.

DR PROSITE; PS00321; RECA_1; 1.

DR PROSITE; PS0162; RECA_2; 1.

DR PROSITE; PS0163; RECA_3; 1.

KW ATP-binding; DNA damage; DNA recombination; DNA repair; DNA-binding;

KW Nucleotide-binding; SOS response.

FT INIT MET 0 By similarity.

FT NP_BIND 66 73 ATP (By similarity).

SQ SEQUENCE 353 AA; 37767 MW; 960F1F2698CAC46B CRC64;

Query Match 91.1%; Score 1528; DB 1; Length 353;

Best Local Similarity 91.8%; Pred. No. 8.5e-94;

Matches 302; Conservative 16; Mismatches 11; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSLDIALGAGGLPMGR 60

DB 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSLDIALGAGGLPMGR 60

QY 61 IVEITYGPSSGKTTLTQVIAAAREGKTCAPDAEHALDPIYARKLGVDIDNLLCSQPD 120

DB 61 IVEITYGPSSGKTTLTQVIAAAREGKTCAPDAEHALDPIYARKLGVDIDNLLCSQPD 120

QY 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGEGIDSHMGLAARWMSQAMRKLKAG 180

DB 121 TGEQALEICDALTSGADVIVVDSVAALTPKAEIEGEGIDSHMGLAARWMSQAMRKLKAG 180

QY 181 NLKOSNTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240

DB 181 NLKNANTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVEGDEVVGS 240

QY 241 ETRVKVVKNTAAAPFKQAEFQILYGEINFGYELVDLGKVEKLEKAGAWTSYNGEKIQG 300

DB 241 ETRVKVVKNTAAAPFKQAEFQILYGEINFGYELVDLGKVEKLEKAGAWTSYNGEKIQG 300

QY 301 GKANATAWLKDNPETAKEIEKKVRELLLS 329

DB 301 GKANSNCYLNKPNKVAALDKKRDMLLS 329

RESULT 15

RECA_SERMA

ID RECA_SERMA

AC P17479;

DT 01-AUG-1990 (Rel. 15, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE RecA protein (Recombinase A).

GN Name=recA;

OS Serratia marcescens.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Serratia.

OX NCBI_taxID=615;

RN [1]

RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].

RC STRAIN=SM6;

RX MEDLINE=90094239; PubMed=2152908;

RA Ball T.K., Wasmuth C.R., Braunagel S.C., Benedik M.J.;

RT "Expression of Serratia marcescens extracellular proteins requires
RL J. Bacteriol. 172:342-349(1990).
CC -1- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the recA family.
CC

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL; M22935; AAA26567.1; -; Genomic_DNA.
CC HSSP; P03017; 2REB.
CC SMR; P17479; 3-328.

CC HAMAP; MF 00268; -; 1.
CC InterPro; IPR003593; AAA ATPase.

CC InterPro; IPR001553; RecA.

CC Pfam; PF00154; RecA; 1.

CC PRINTS; PR00142; RECA.

CC ProDom; PD000229; RecA; 1.

CC SMART; SM00382; AAA; 1.

CC TIGRFAMs; TIGR02012; tigrfam_recA; 1.

CC PROSITE; PS00321; RECA_1; 1.

CC PROSITE; PS0162; RECA_2; 1.

CC PROSITE; PS0163; RECA_3; 1.

KW ATP-binding; DNA damage; DNA recombination; DNA repair; DNA-binding;

KW Nucleotide-binding; SOS response.

FT INIT MET 0 By similarity.

FT NP_BIND 66 73 ATP (By similarity).

SQ SEQUENCE 353 AA; 37778 MW; A97369360970F814 CRC64;

Query Match 90.8%; Score 1523; DB 1; Length 353;

Best Local Similarity 92.1%; Pred. No. 1.8e-93;

Matches 302; Conservative 14; Mismatches 12; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSLDIALGAGGLPMGR 60

DB 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSLDIALGAGGLPMGR 60

QY 61 IVEITYGPSSGKTTLTQVIAAAREGKTCAPDAEHALDPIYARKLGVDIDNLLCSQPD 120

DB 61 IVEITYGPSSGKTTLTQVIAAAREGKTCAPDAEHALDPIYARKLGVDIDNLLCSQPD 120

QY 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGEGIDSHMGLAARWMSQAMRKLKAG 180

DB 121 TGEQALEICDALTSGADVIVVDSVAALTPKAEIEGEGIDSHMGLAARWMSQAMRKLKAG 180

QY 181 NLKOSNTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240

DB 181 NLKNANTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVEGDEVVGS 240

QY 241 ETRVKVVKNTAAAPFKQAEFQILYGEINFGYELVDLGKVEKLEKAGAWTSYNGEKIQG 300

DB 241 ETRVKVVKNTAAAPFKQAEFQILYGEINFGYELVDLGKVEKLEKAGAWTSYNGEKIQG 300

QY 301 GKANATAWLKDNPETAKEIEKKVRELLLS 328

DB 301 GKANACNFKLENPAIAAELDKLRDLLL 328

Search completed: February 16, 2006, 01:24:01

Job time : 153 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 01:24:19 ; Search time 32.5 Seconds
(without alignments)
852.195 Million cell updates/sec

Title: US-10-733-782-3
Perfect score: 1678
Sequence: 1 AIDENKQALAAALGQIEKQ.....AKETKVKRELLSNPNSTP 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1674	99.8	353	1	US-08-097-831-6
2	1674	99.8	358	2	US-09-626-410-7
3	1674	99.8	358	2	US-09-626-410-9
4	1674	99.8	358	2	US-09-626-410-14
5	1674	99.8	358	2	US-09-116-188-7
6	1674	99.8	358	2	US-09-116-188-9
7	1674	99.8	358	2	US-09-116-188-14
8	1674	99.8	358	2	US-09-626-047-7
9	1674	99.8	358	2	US-09-626-047-9
10	1674	99.8	358	2	US-09-626-047-14
11	1674	99.8	358	2	US-09-626-343-7
12	1674	99.8	358	2	US-09-626-343-9
13	1674	99.8	358	2	US-09-626-343-14
14	1674	99.8	358	2	US-09-354-922-8
15	1674	99.8	358	2	US-09-354-922-10
16	1674	99.8	358	2	US-09-354-922-15
17	1674	99.8	358	2	US-09-516-051-7
18	1674	99.8	358	2	US-09-516-051-9
19	1674	99.8	358	2	US-09-516-051-14
20	1674	99.8	358	2	US-09-516-695B-8
21	1674	99.8	358	2	US-09-516-695B-10
22	1674	99.8	358	2	US-09-516-695B-15
23	1669	99.5	358	2	US-09-626-410-8
24	1669	99.5	358	2	US-09-626-410-11
25	1669	99.5	358	2	US-09-116-188-8
26	1669	99.5	358	2	US-09-116-188-11
27	1669	99.5	358	2	US-09-626-047-8

28	1669	99.5	358	2	US-09-626-047-11	Sequence 11, Appl
29	1669	99.5	358	2	US-09-626-343-8	Sequence 8, Appl
30	1669	99.5	358	2	US-09-626-343-11	Sequence 11, Appl
31	1669	99.5	358	2	US-09-354-922-9	Sequence 9, Appl
32	1669	99.5	358	2	US-09-354-922-12	Sequence 12, Appl
33	1669	99.5	358	2	US-09-516-051-8	Sequence 8, Appl
34	1669	99.5	358	2	US-09-516-051-11	Sequence 11, Appl
35	1669	99.5	358	2	US-09-516-695B-9	Sequence 9, Appl
36	1669	99.5	358	2	US-09-516-695B-12	Sequence 12, Appl
37	1656	98.7	358	2	US-09-626-410-12	Sequence 12, Appl
38	1656	98.7	358	2	US-09-116-188-12	Sequence 12, Appl
39	1656	98.7	358	2	US-09-626-047-12	Sequence 12, Appl
40	1656	98.7	358	2	US-09-626-343-12	Sequence 12, Appl
41	1656	98.7	358	2	US-09-354-922-13	Sequence 13, Appl
42	1656	98.7	358	2	US-09-516-051-12	Sequence 12, Appl
43	1656	98.7	358	2	US-09-516-695B-13	Sequence 13, Appl
44	1655	98.6	358	2	US-09-626-410-10	Sequence 10, Appl
45	1655	98.6	358	2	US-09-116-188-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-097-831-6
; Sequence 6, Application US/08097831
; Patent No. 5510473
; GENERAL INFORMATION:
; APPLICANT: Camerini-Otero, Rafael D.
; APPLICANT: Angov, Evangelina
; TITLE OF INVENTION: Cloning and Expression of Taq recA
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097,831
; FILING DATE: 19930726
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E.
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: NIH066.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
; TELEFAX: 714-760-9502
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; US-08-097-831-6

Query Match 99.8%; Score 1674; DB 1; Length 353;
Best Local Similarity 99.7%; Pred. No. 3.7e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQALAAALGQIEKQFGKSGMRGDRSMVKTISTGSLSLDIALGAGGLPMGR 60
|||||

Db 2 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDVETISTGSLSLDIALGAGGLPMGR 61
Qy 61 IVEIYGPSSSGKTTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDIDNLLCSQPD 120
Db 62 IVEIYGPSSSGKTTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDIDNLLCSQPD 121
Qy 121 TGEQALEICDALARSGAVDVIVDVSVAALTPKAEIEGEIGDGHMGLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALARSGAVDVIVDVSVAALTPKAEIEGEIGDGHMGLAARMMSQAMRKLKAG 181
Qy 181 NLKQNTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 182 NLKQNTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 241
Qy 241 ETRVVKVKNKIAAPPKQAEFOILYGEINFGYELVDLGKVEKLEKAGAWTSYKGEKIGQ 300
Db 242 ETRVVKVKNKIAAPPKQAEFOILYGEINFGYELVDLGKVEKLEKAGAWTSYKGEKIGQ 301
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 302 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 336

RESULT 2

US-09-626-410-7
; Sequence 7, Application US/09626410
; Patent No. 6287862
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/626,410
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-626-410-7

Query Match 99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDVETISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDVETISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPSSSGKTTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDIDNLLCSQPD 120
Db 67 IVEIYGPSSSGKTTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDIDNLLCSQPD 126
Qy 121 TGEQALEICDALARSGAVDVIVDVSVAALTPKAEIEGEIGDGHMGLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGAVDVIVDVSVAALTPKAEIEGEIGDGHMGLAARMMSQAMRKLKAG 186
Qy 181 NLKQNTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKQNTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246

Qy 241 ETRVVKVKNKIAAPPKQAEFOILYGEINFGYELVDLGKVEKLEKAGAWTSYKGEKIGQ 300
Db 247 ETRVVKVKNKIAAPPKQAEFOILYGEINFGYELVDLGKVEKLEKAGAWTSYKGEKIGQ 306
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 3

US-09-626-410-9
; Sequence 9, Application US/09626410
; Patent No. 6287862
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/626,410
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-626-410-9

Query Match 99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDVETISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDVETISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPSSSGKTTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDIDNLLCSQPD 120
Db 67 IVEIYGPSSSGKTTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDIDNLLCSQPD 126
Qy 121 TGEQALEICDALARSGAVDVIVDVSVAALTPKAEIEGEIGDGHMGLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGAVDVIVDVSVAALTPKAEIEGEIGDGHMGLAARMMSQAMRKLKAG 186
Qy 181 NLKQNTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKQNTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
Qy 241 ETRVVKVKNKIAAPPKQAEFOILYGEINFGYELVDLGKVEKLEKAGAWTSYKGEKIGQ 300
Db 247 ETRVVKVKNKIAAPPKQAEFOILYGEINFGYELVDLGKVEKLEKAGAWTSYKGEKIGQ 306
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 4

US-09-626-410-14
; Sequence 14, Application US/09626410
; Patent No. 6287862
; GENERAL INFORMATION:

APPLICANT: DEL CARDAYRE, STEPHEN
APPLICANT: TOBIN, MATTHEW
APPLICANT: STEMMER, WILLEM P.C.
APPLICANT: NESS, JON E.
APPLICANT: MINSHULL, JEREMY
APPLICANT: PATTEN, PHILLIP
APPLICANT: SUBRAMANIA, VENKITSWARAN
APPLICANT: CASTLE, LINDA
APPLICANT: KREBBER, CLAUD M.
APPLICANT: BASS, STEVE

TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE

FILE REFERENCE: 02-020720US

CURRENT APPLICATION NUMBER: US/09/626,410

CURRENT FILING DATE: 2000-07-26

PRIOR APPLICATION NUMBER: 09/116,188

PRIOR FILING DATE: 07-15-1998

NUMBER OF SEQ ID NOS: 14

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 14

LENGTH: 358

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: consensus

OTHER INFORMATION: e. coli sequence

US-09-626-410-14

Query Match 99.8%; Score 1674; DB 2; Length 358;

Best Local Similarity 99.7%; Pred. No. 3.8e-167;

Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60

Db 7 AIDENKQKALAAALGQIEKQFGKSGIMRLGDRSDMDVKTISTGSLSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTLTQVIAAAREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120

Db 67 IVEIYGPESSGKTTLTQVIAAAREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 126

Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGSHMGLAARMMSQAMRKLKLAG 180

Db 127 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGSHMGLAARMMSQAMRKLKLAG 186

Qy 181 NLKQSNLTLLIIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGNNVVG 240

Db 187 NLKQSNLTLLIIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGNNVVG 246

Qy 241 ETRVVKVKNKIAAPFKQAEFQILYGEINFGELVDLGKVKLEKAGAWYSYKGEKIGQ 300

Db 247 ETRVVKVKNKIAAPFKQAEFQILYGEINFGELVDLGKVKLEKAGAWYSYKGEKIGQ 306

Qy 301 GKANATAWKDNPTAKEIEKKVRELLLSNPNSTP 335

Db 307 GKANATAWKDNPTAKEIEKKVRELLLSNPNSTP 341

RESULT 5

US-09-116-188-7

Sequence 7, Application US/09116188

Patent No. 6326204

GENERAL INFORMATION:

APPLICANT: DEL CARDAYRE, STEPHEN

APPLICANT: TOBIN, MATTHEW

APPLICANT: STEMMER, WILLEM P.C.

APPLICANT: NESS, JON E.

APPLICANT: MINSHULL, JEREMY

APPLICANT: PATTEN, PHILLIP

APPLICANT: SUBRAMANIA, VENKITSWARAN

APPLICANT: CASTLE, LINDA

APPLICANT: KREBBER, CLAUD M.

APPLICANT: BASS, STEVE

TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE

TITLE OF INVENTION: SEQUENCE RECOMBINATION

FILE REFERENCE: 02-020720US

CURRENT APPLICATION NUMBER: US/09/116,188

CURRENT FILING DATE: 1998-07-15

NUMBER OF SEQ ID NOS: 14

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 7

LENGTH: 358

TYPE: PRT

ORGANISM: Escherichia coli

US-09-116-188-7

Query Match 99.8%; Score 1674; DB 2; Length 358;

Best Local Similarity 99.7%; Pred. No. 3.8e-167;

Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60

Db 7 AIDENKQKALAAALGQIEKQFGKSGIMRLGDRSDMDVKTISTGSLSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTLTQVIAAAREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120

Db 67 IVEIYGPESSGKTTLTQVIAAAREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 126

Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGSHMGLAARMMSQAMRKLKLAG 180

Db 127 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGSHMGLAARMMSQAMRKLKLAG 186

Qy 181 NLKQSNLTLLIIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGNNVVG 240

Db 187 NLKQSNLTLLIIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGNNVVG 246

Qy 241 ETRVVKVKNKIAAPFKQAEFQILYGEINFGELVDLGKVKLEKAGAWYSYKGEKIGQ 300

Db 247 ETRVVKVKNKIAAPFKQAEFQILYGEINFGELVDLGKVKLEKAGAWYSYKGEKIGQ 306

Qy 301 GKANATAWKDNPTAKEIEKKVRELLLSNPNSTP 335

Db 307 GKANATAWKDNPTAKEIEKKVRELLLSNPNSTP 341

RESULT 6

US-09-116-188-9

Sequence 9, Application US/09116188

Patent No. 6326204

GENERAL INFORMATION:

APPLICANT: DEL CARDAYRE, STEPHEN

APPLICANT: TOBIN, MATTHEW

APPLICANT: STEMMER, WILLEM P.C.

APPLICANT: NESS, JON E.

APPLICANT: MINSHULL, JEREMY

APPLICANT: PATTEN, PHILLIP

APPLICANT: SUBRAMANIA, VENKITSWARAN

APPLICANT: CASTLE, LINDA

APPLICANT: KREBBER, CLAUD M.

APPLICANT: BASS, STEVE

TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE

FILE REFERENCE: 02-020720US

CURRENT APPLICATION NUMBER: US/09/116,188

CURRENT FILING DATE: 1998-07-15

NUMBER OF SEQ ID NOS: 14

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 9

LENGTH: 358

TYPE: PRT

ORGANISM: Escherichia coli

US-09-116-188-9

Query Match 99.8%; Score 1674; DB 2; Length 358;

Best Local Similarity 99.7%; Pred. No. 3.8e-167;

Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLDIALGAGGLPMGR 66
QY 61 IVEIYGPRESSGKTTTLQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPRESSGKTTTLQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 186
QY 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVGS 240
DB 187 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVGS 246
QY 241 ETRVVKVKNKTAAPFKQAEFQILYEGEINFYGGELVDLGKVKELIEKAGAWSYKGEKIGQ 300
DB 247 ETRVVKVKNKTAAPFKQAEFQILYEGEINFYGGELVDLGKVKELIEKAGAWSYKGEKIGQ 306
QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
DB 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 7

US-09-116-188-14
; Sequence 14, Application US/09116188
; Patent No. 6326204
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P. C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/116,188
; CURRENT FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
US-09-116-188-14

Query Match 99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLDIALGAGGLPMGR 66
QY 61 IVEIYGPRESSGKTTTLQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPRESSGKTTTLQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 186
QY 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVGS 240

DB 187 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVGS 246
QY 241 ETRVVKVKNKTAAPFKQAEFQILYEGEINFYGGELVDLGKVKELIEKAGAWSYKGEKIGQ 300
DB 247 ETRVVKVKNKTAAPFKQAEFQILYEGEINFYGGELVDLGKVKELIEKAGAWSYKGEKIGQ 306
QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
DB 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 8

US-09-626-047-7
; Sequence 7, Application US/09626047
; Patent No. 6335198
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P. C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/626,047
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-626-047-7

Query Match 99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLDIALGAGGLPMGR 66
QY 61 IVEIYGPRESSGKTTTLQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPRESSGKTTTLQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 186
QY 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVGS 240
DB 187 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVGS 246
QY 241 ETRVVKVKNKTAAPFKQAEFQILYEGEINFYGGELVDLGKVKELIEKAGAWSYKGEKIGQ 300
DB 247 ETRVVKVKNKTAAPFKQAEFQILYEGEINFYGGELVDLGKVKELIEKAGAWSYKGEKIGQ 306
QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
DB 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 9

US-09-626-047-9

; Sequence 9, Application US/09626047
; Patent No. 6335198
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: SEQUENCE RECOMBINATION
; CURRENT APPLICATION NUMBER: US/09/626,047
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-626-047-9

Query Match 99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGDRSDMDVKTISTGSLDIALGAGGLPMGR 60
Db 7 AIDENKOKALAAALGQIEKQFGKGSIMRLGDRSDMDVKTISTGSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126

Qy 121 TGEQALEICDALARSGADVIVDSVAALTPKAEIEGISHMGLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGADVIVDSVAALTPKAEIEGISHMGLAARMMSQAMRKLKAG 186

Qy 181 NLKQSTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGNNVVS 240
Db 187 NLKQSTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGNNVVS 246

Qy 241 ETRVKKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLIKAGAWTSYKGEKIGQ 300
Db 247 ETRVKKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLIKAGAWTSYKGEKIGQ 306

Qy 301 GKANATAWKDNPETAKEIEKKVRELLSNPNSTP 335
Db 307 GKANATAWKDNPETAKEIEKKVRELLSNPNSTP 341

RESULT 10
US-09-626-047-14
; Sequence 14, Application US/09626047
; Patent No. 6335198
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE

; TITLE OF INVENTION: SEQUENCE RECOMBINATION
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/626,047
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; OTHER INFORMATION: e. coli sequence
US-09-626-047-14

Query Match 99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGDRSDMDVKTISTGSLDIALGAGGLPMGR 60
Db 7 AIDENKOKALAAALGQIEKQFGKGSIMRLGDRSDMDVKTISTGSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126

Qy 121 TGEQALEICDALARSGADVIVDSVAALTPKAEIEGISHMGLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGADVIVDSVAALTPKAEIEGISHMGLAARMMSQAMRKLKAG 186

Qy 181 NLKQSTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGNNVVS 240
Db 187 NLKQSTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGNNVVS 246

Qy 241 ETRVKKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLIKAGAWTSYKGEKIGQ 300
Db 247 ETRVKKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLIKAGAWTSYKGEKIGQ 306

Qy 301 GKANATAWKDNPETAKEIEKKVRELLSNPNSTP 335
Db 307 GKANATAWKDNPETAKEIEKKVRELLSNPNSTP 341

RESULT 11
US-09-626-343-7
; Sequence 7, Application US/09626343
; Patent No. 6352859
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW P.C.
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020727US
; CURRENT APPLICATION NUMBER: US/09/626,343
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 358
; TYPE: PRT

```

; ORGANISM: Escherichia coli
US-09-626-343-7

Query Match          99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSDMDVKTISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPESSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 126
Qy 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 186
Qy 181 NLKQNTLLIFINQIRMKIGVMFGNPGPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKQNTLLIFINQIRMKIGVMFGNPGPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
Qy 241 ETRVWVKNKIAAPFKQAEFOILYGEINFGYELVDLGKVEKLTIEKAGAWTSYKGEKIGQ 300
Db 247 ETRVWVKNKIAAPFKQAEFOILYGEINFGYELVDLGKVEKLTIEKAGAWTSYKGEKIGQ 306
Qy 301 GKANATAWLKONPETAKIEKKVRELLLSNPSTP 335
Db 307 GKANATAWLKONPETAKIEKKVRELLLSNPSTP 341

RESULT 12
US-09-626-343-9
; Sequence 9, Application US/09626343
; Patent No. 6352859
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P. C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020727US
; CURRENT APPLICATION NUMBER: US/09/626,343
; CURRENT FILING DATE: 2000-07-26
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-626-343-9

Query Match          99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSDMDVKTISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPESSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 126
Qy 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 186
Qy 181 NLKQNTLLIFINQIRMKIGVMFGNPGPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKQNTLLIFINQIRMKIGVMFGNPGPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
Qy 241 ETRVWVKNKIAAPFKQAEFOILYGEINFGYELVDLGKVEKLTIEKAGAWTSYKGEKIGQ 300
Db 247 ETRVWVKNKIAAPFKQAEFOILYGEINFGYELVDLGKVEKLTIEKAGAWTSYKGEKIGQ 306
Qy 301 GKANATAWLKONPETAKIEKKVRELLLSNPSTP 335
Db 307 GKANATAWLKONPETAKIEKKVRELLLSNPSTP 341

RESULT 13
US-09-626-343-14
; Sequence 14, Application US/09626343
; Patent No. 6352859
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P. C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020727US
; CURRENT APPLICATION NUMBER: US/09/626,343
; CURRENT FILING DATE: 2000-07-26
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; OTHER INFORMATION: e. coli sequence
US-09-626-343-14

Query Match          99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSDMDVKTISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPESSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 126
Qy 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 186
Qy 181 NLKQNTLLIFINQIRMKIGVMFGNPGPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKQNTLLIFINQIRMKIGVMFGNPGPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
Qy 241 ETRVWVKNKIAAPFKQAEFOILYGEINFGYELVDLGKVEKLTIEKAGAWTSYKGEKIGQ 300
Db 247 ETRVWVKNKIAAPFKQAEFOILYGEINFGYELVDLGKVEKLTIEKAGAWTSYKGEKIGQ 306
```

Db 247 ETRVVKVKNKIAAPFKQAEFOILYGEINFGELVDLGVKEKLEKAGAWTSYKGEKIGQ 306
Qy 301 GKANATAWLKNPETAKEIEKKVRELLSNPNSTP 335
Db 307 GKANATAWLKNPETAKEIEKKVRELLSNPNSTP 341

RESULT 14

US-09-354-922-8
; Sequence 8, Application US/09354922
; Patent No. 6379964

GENERAL INFORMATION:

; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.

; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; TITLE OF INVENTION: SEQUENCE RECOMBINATION
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/09/354.922
; CURRENT FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 8

; LENGTH: 358

; TYPE: PRT

; ORGANISM: Escherichia coli

US-09-354-922-8

Query Match 99.8%; Score 1674; DB 2; Length 358;

Best Local Similarity 99.7%; Pred. No. 3.8e-167;

Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGOIEKQFGKGSIMRLGDRSDMDVKTISTGSLDIALGAGGLPMGR 60

Db 7 AIDENKOKALAAALGOIEKQFGKGSIMRLGDRSDMDVKTISTGSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120

Db 67 IVEIYGPESSGKTTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126

Qy 121 TGEQALBICDALARSGAVDVIVDVSVAALTPKABIEGIEGDSHMGGLAARMMSQAMRKLKAG 180

Db 127 TGEQALBICDALARSGAVDVIVDVSVAALTPKABIEGIEGDSHMGGLAARMMSQAMRKLKAG 186

Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240

Db 187 NLKQSNLTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246

Qy 241 ETRVVKVKNKIAAPFKQAEFOILYGEINFGELVDLGVKEKLEKAGAWTSYKGEKIGQ 300

Db 247 ETRVVKVKNKIAAPFKQAEFOILYGEINFGELVDLGVKEKLEKAGAWTSYKGEKIGQ 306

RESULT 15

US-09-354-922-10

; Sequence 10, Application US/09354922

; Patent No. 6379964

GENERAL INFORMATION:

; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.

; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; TITLE OF INVENTION: SEQUENCE RECOMBINATION
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/09/354.922
; CURRENT FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 10

; LENGTH: 358

; TYPE: PRT

; ORGANISM: Escherichia coli

US-09-354-922-10

Query Match 99.8%; Score 1674; DB 2; Length 358;

Best Local Similarity 99.7%; Pred. No. 3.8e-167;

Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGOIEKQFGKGSIMRLGDRSDMDVKTISTGSLDIALGAGGLPMGR 60

Db 7 AIDENKOKALAAALGOIEKQFGKGSIMRLGDRSDMDVKTISTGSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120

Db 67 IVEIYGPESSGKTTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126

Qy 121 TGEQALBICDALARSGAVDVIVDVSVAALTPKABIEGIEGDSHMGGLAARMMSQAMRKLKAG 180

Db 127 TGEQALBICDALARSGAVDVIVDVSVAALTPKABIEGIEGDSHMGGLAARMMSQAMRKLKAG 186

Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240

Db 187 NLKQSNLTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246

Qy 241 ETRVVKVKNKIAAPFKQAEFOILYGEINFGELVDLGVKEKLEKAGAWTSYKGEKIGQ 300

Db 247 ETRVVKVKNKIAAPFKQAEFOILYGEINFGELVDLGVKEKLEKAGAWTSYKGEKIGQ 306

Search completed: February 16, 2006, 01:26:08

Job time : 32.5 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 01:39:50 ; Search time 123 Seconds
(without alignments)
1137.990 Million cell updates/sec

Title: US-10-733-782-3
Perfect score: 1678
Sequence: 1 AIDENKQKALAAALGQIEKQ.....AKEIEKKVRELLSNPNSTP 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA_Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1678	100.0	335	4	US-10-733-782-3
2	1674	99.8	335	4	US-10-733-782-1
3	1674	99.8	352	4	US-10-733-782-5
4	1674	99.8	353	4	US-10-282-122A-43319
5	1674	99.8	358	4	US-10-194-686-8
6	1674	99.8	358	4	US-10-194-686-10
7	1674	99.8	358	4	US-10-194-686-15
8	1669	99.5	358	4	US-10-194-686-9
9	1669	99.5	358	4	US-10-194-686-12
10	1656	98.7	358	4	US-10-194-686-13
11	1655	98.6	358	4	US-10-194-686-11
12	1637	97.6	353	4	US-10-282-122A-75948
13	1609	95.9	352	4	US-10-282-122A-59525
14	1589	94.7	352	4	US-10-282-122A-55632
15	1530	91.2	356	4	US-10-282-122A-78524
16	1500	89.4	355	4	US-10-282-122A-68593
17	1463	87.2	351	4	US-10-282-122A-72895
18	1432	85.3	412	4	US-10-282-122A-77075
19	1349	80.4	354	4	US-10-282-122A-67415
20	1328	79.1	354	4	US-10-282-122A-58235
21	1268	75.6	346	4	US-10-282-122A-66505
22	1257	74.9	355	4	US-10-282-122A-67942
23	1255	74.8	349	4	US-10-282-122A-44768
24	1255	74.8	369	4	US-10-282-122A-69802
25	1242.5	74.0	349	4	US-10-282-122A-63392
26	1238.5	73.8	342	4	US-10-751-928-20
27	1237.5	73.7	348	4	US-10-282-122A-61049

28	1213.5	72.3	348	5	US-10-988-943-33	Sequence 33, Appl
29	1212.5	72.3	348	4	US-10-282-122A-65975	Sequence 65975, A
30	1208.5	72.0	375	4	US-10-282-122A-65031	Sequence 65031, A
31	1200.5	71.5	356	4	US-10-282-122A-47848	Sequence 47848, A
32	1200.5	71.5	356	4	US-10-282-122A-50519	Sequence 50519, A
33	1196.5	71.3	358	4	US-10-282-122A-49452	Sequence 49452, A
34	1180.5	70.4	353	4	US-10-282-122A-51342	Sequence 51342, A
35	1148	68.4	377	4	US-10-156-761-10029	Sequence 10029, A
36	1139	67.9	356	4	US-10-282-122A-52244	Sequence 52244, A
37	1135	67.6	354	4	US-10-282-122A-53057	Sequence 53057, A
38	1131	67.4	349	4	US-10-282-122A-51953	Sequence 51953, A
39	1112	66.3	350	4	US-10-282-122A-61898	Sequence 61898, A
40	1107	66.0	350	4	US-10-203-927A-2	Sequence 2, Appli
41	1107	66.0	356	5	US-10-501-282-5628	Sequence 5628, Ap
42	1107	66.0	361	5	US-10-501-282-5630	Sequence 5630, Ap
43	1095	65.3	348	4	US-10-282-122A-71634	Sequence 71634, A
44	1089	64.9	347	4	US-10-282-122A-70318	Sequence 70318, A
45	1089	64.9	357	4	US-10-724-972A-4382	Sequence 4382, Ap

ALIGNMENTS

RESULT 1
US-10-733-782-3
; Sequence 3, Application US/10733782
; Publication No. US20040157248A1
; GENERAL INFORMATION:
; APPLICANT: Cox, Michael
; APPLICANT: Lusetti, Shelley
; APPLICANT: Eggleter, Aimee
; TITLE OF INVENTION: RecA Mutants
; FILE REFERENCE: 960296.99501
; CURRENT APPLICATION NUMBER: US/10733,782
; CURRENT FILING DATE: 2003-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 3
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-733-782-3

Query Match	100.0%	Score 1678;	DB 4;	Length 335;
Best Local Similarity	100.0%	Pred. No. 2.7e-152;		
Matches	335;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDKTI	STGSLSLDIALGAGGLPMGR	60
Db	1	AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDKTI	STGSLSLDIALGAGGLPMGR	60
Qy	61	IVEIYGPSSGKTTTLQVIAAAREKTCAPIDAHAALDPIYARKLGV	IDNLLCSQPD	120
Db	61	IVEIYGPSSGKTTTLQVIAAAREKTCAPIDAHAALDPIYARKLGV	IDNLLCSQPD	120
Qy	121	TGEALRICDALARSGAVDVIVDVSVAALTPKAEIEGIDSHMGLAAR	MMSQMRKLAG	180
Db	121	TGEALRICDALARSGAVDVIVDVSVAALTPKAEIEGIDSHMGLAAR	MMSQMRKLAG	180
Qy	181	NLKQSNLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRR	IGAVKEGNNVGS	240
Db	181	NLKQSNLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRR	IGAVKEGNNVGS	240
Qy	241	ETRVKVVKNKIAAPFKQAEFQILYGEINPFYGLVDLGVEKLEKAG	AWYSYKGEKIQ	300
Db	241	ETRVKVVKNKIAAPFKQAEFQILYGEINPFYGLVDLGVEKLEKAG	AWYSYKGEKIQ	300
Qy	301	GKANATAWLNDONPETAKEIEKKVRELLSNPNSTP	335	
Db	301	GKANATAWLNDONPETAKEIEKKVRELLSNPNSTP	335	

RESULT 2

```
US-10-733-782-1
; Sequence 1, Application US/107333782
; Publication No. US20040157248A1
; GENERAL INFORMATION:
; APPLICANT: Cox, Michael
; APPLICANT: Lusettti, Shelley
; APPLICANT: Egglar, Aimee
; TITLE OF INVENTION: RecA Mutants
; FILE REFERENCE: 960296.99501
; CURRENT APPLICATION NUMBER: US/10/733,782
; CURRENT FILING DATE: 2003-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-733-782-1
Query Match 99.8%; Score 1674; DB 4; Length 335;
Best Local Similarity 99.7%; Pred. No. 6.5e-152;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 60
Db 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 60
Qy 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLKAG 180
Db 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLKAG 180
Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Qy 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKLEKAGAWYSYKGEKIGQ 300
Db 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKLEKAGAWYSYKGEKIGQ 300
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335

RESULT 3
US-10-733-782-5
; Sequence 5, Application US/107333782
; Publication No. US20040157248A1
; GENERAL INFORMATION:
; APPLICANT: Cox, Michael
; APPLICANT: Lusettti, Shelley
; APPLICANT: Egglar, Aimee
; TITLE OF INVENTION: RecA Mutants
; FILE REFERENCE: 960296.99501
; CURRENT APPLICATION NUMBER: US/10/733,782
; CURRENT FILING DATE: 2003-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-733-782-5
Query Match 99.8%; Score 1674; DB 4; Length 352;
Best Local Similarity 99.7%; Pred. No. 7e-152;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 60
Db 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 60
Qy 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLKAG 180
Db 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLKAG 180
Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Qy 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKLEKAGAWYSYKGEKIGQ 300
Db 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKLEKAGAWYSYKGEKIGQ 300
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335

RESULT 4
US-10-282-122A-43319
; Sequence 43319, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43319
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-282-122A-43319
```

Query Match 99.8%; Score 1674; DB 4; Length 353;
Best Local Similarity 99.7%; Pred. No. 7e-152;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 61
Qy 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 121
Qy 121 TGEQALEICDALARSGAVDVIVVDSVAALTPKABIEGIEGDSHGLAARMMSQAMRKLAG 180
Db 122 TGEQALEICDALARSGAVDVIVVDSVAALTPKABIEGIEGDSHGLAARMMSQAMRKLAG 181
Qy 181 NLKQSNLTLLIFINQIRMKIGVMFNGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
Db 182 NLKQSNLTLLIFINQIRMKIGVMFNGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 241
Qy 241 ETRVKKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVEKLEKAGAWTSYKGEKIGQ 300
Db 242 ETRVKKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVEKLEKAGAWTSYKGEKIGQ 301
Qy 301 GKANATAWLNDKONPETAKEIEKKVRELLSNPNSTP 335
Db 302 GKANATAWLNDKONPETAKEIEKKVRELLSNPNSTP 336

RESULT 5

US-10-194-686-8
; Sequence 8, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-194-686-8

Query Match 99.8%; Score 1674; DB 4; Length 358;
Best Local Similarity 99.7%; Pred. No. 7.2e-152;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 60
Db 7 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126
Qy 121 TGEQALEICDALARSGAVDVIVVDSVAALTPKABIEGIEGDSHGLAARMMSQAMRKLAG 180
Db 127 TGEQALEICDALARSGAVDVIVVDSVAALTPKABIEGIEGDSHGLAARMMSQAMRKLAG 186
Qy 181 NLKQSNLTLLIFINQIRMKIGVMFNGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
Db 187 NLKQSNLTLLIFINQIRMKIGVMFNGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 246
Qy 241 ETRVKKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVEKLEKAGAWTSYKGEKIGQ 300
Db 247 ETRVKKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVEKLEKAGAWTSYKGEKIGQ 306
Qy 301 GKANATAWLNDKONPETAKEIEKKVRELLSNPNSTP 335
Db 307 GKANATAWLNDKONPETAKEIEKKVRELLSNPNSTP 341

RESULT 6

US-10-194-686-10
; Sequence 10, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-194-686-10

Query Match 99.8%; Score 1674; DB 4; Length 358;
Best Local Similarity 99.7%; Pred. No. 7.2e-152;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 60
Db 7 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 66
Qy 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126
Qy 121 TGEQALEICDALARSGAVDVIVVDSVAALTPKABIEGIEGDSHGLAARMMSQAMRKLAG 180
Db 127 TGEQALEICDALARSGAVDVIVVDSVAALTPKABIEGIEGDSHGLAARMMSQAMRKLAG 186
Qy 181 NLKQSNLTLLIFINQIRMKIGVMFNGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240

```
||||| 187 NLKQSNLTLLIFINQIRMKIGVFNPGPNTTTCGNALKFYASVRLDIRRIGAVKEGENVVGS 246
||||| 241 ETRVKKVKNKTAAPKQAEFQILYEGEINFGYELVDLGVKEKLIBKAGAWTSYKGEKIGQ 300
||||| 247 ETRVKKVKNKTAAPKQAEFQILYEGEINFGYELVDLGVKEKLIBKAGAWTSYKGEKIGQ 306

QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 335
||||| 307 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 341

RESULT 7
US-10-194-686-15
; Sequence 15, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; TITLE OF INVENTION: SEQUENCE RECOMBINATION
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; PRIOR FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; OTHER INFORMATION: e. coli sequence
US-10-194-686-15

Query Match 99.8%; Score 1674; DB 4; Length 358;
Best Local Similarity 99.7%; Pred. No. 7.2e-152;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60
||||| 7 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSDMDVKTISTGSLSLDIALGAGGLPMGR 66

QY 61 IVEIYGPSSGKTTTLQVIAAAQREGTKCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
||||| 67 IVEIYGPSSGKTTTLQVIAAAQREGTKCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126

QY 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGEGIDSHMGLAARMMSQAMRKLAG 180
||||| 127 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGEGIDSHMGLAARMMSQAMRKLAG 186

QY 181 NLKQSNLTLLIFINQIRMKIGVFNPGPNTTTCGNALKFYASVRLDIRRIGAVKEGENVVGS 240
||||| 187 NLKQSNLTLLIFINQIRMKIGVFNPGPNTTTCGNALKFYASVRLDIRRIGAVKEGENVVGS 246

QY 241 ETRVKKVKNKTAAPKQAEFQILYEGEINFGYELVDLGVKEKLIBKAGAWTSYKGEKIGQ 300
||||| 247 ETRVKKVKNKTAAPKQAEFQILYEGEINFGYELVDLGVKEKLIBKAGAWTSYKGEKIGQ 306

Db 187 NLKQSNLTLLIFINQIRMKIGVFNPGPNTTTCGNALKFYASVRLDIRRIGAVKEGENVVGS 246
||||| 241 ETRVKKVKNKTAAPKQAEFQILYEGEINFGYELVDLGVKEKLIBKAGAWTSYKGEKIGQ 300
||||| 247 ETRVKKVKNKTAAPKQAEFQILYEGEINFGYELVDLGVKEKLIBKAGAWTSYKGEKIGQ 306
```

```
QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 335
||||| 307 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 341

RESULT 8
US-10-194-686-9
; Sequence 9, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; TITLE OF INVENTION: SEQUENCE RECOMBINATION
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; PRIOR FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-194-686-9

Query Match 99.5%; Score 1669; DB 4; Length 358;
Best Local Similarity 99.1%; Pred. No. 2.2e-151;
Matches 332; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60
||||| 7 AIDENKOKALATALGQIEKQFGKGSIMRLGEDRSDMDVKTISTGSLSLDIALGAGGLPMGR 66

QY 61 IVEIYGPSSGKTTTLQVIAAAQREGTKCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
||||| 67 IVEIYGPSSGKTTTLQVIAAAQREGTKCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126

QY 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGEGIDSHMGLAARMMSQAMRKLAG 180
||||| 127 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGEGIDSHMGLAARMMSQAMRKLAG 186

QY 181 NLKQSNLTLLIFINQIRMKIGVFNPGPNTTTCGNALKFYASVRLDIRRIGAVKEGENVVGS 240
||||| 187 NLKQSNLTLLIFINQIRMKIGVFNPGPNTTTCGNALKFYASVRLDIRRIGAVKEGENVVGS 246

QY 241 ETRVKKVKNKTAAPKQAEFQILYEGEINFGYELVDLGVKEKLIBKAGAWTSYKGEKIGQ 300
||||| 247 ETRVKKVKNKTAAPKQAEFQILYEGEINFGYELVDLGVKEKLIBKAGAWTSYKGEKIGQ 306

QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 335
||||| 307 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 341

RESULT 9
US-10-194-686-12
; Sequence 12, Application US/10194686
```



```

; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-194-686-12

Query Match          99.5%; Score 1669; DB 4; Length 358;
Best Local Similarity 99.1%; Pred. No. 2.2e-151;
Matches 332; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGOIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 60
Db 7 AIDENKOKALAAALGOIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126

Qy 121 TGEQALBICDALARSGAVDIVVDSVAALTPKABIEIEIGDGHMGLAARMMSQAMRKLKAG 180
Db 127 TGEQALBICDALARSGAVDIVVDSVAALTPKABIEIEIGDGHMGLAARMMSQAMRKLKAG 186

Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246

Qy 241 ETRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLEKAGAWYSYKGEKIQ 300
Db 247 ETRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLEKAGAWYSYKGEKIQ 306

Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 10
US-10-194-686-13
; Sequence 13, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE

```

```

; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-194-686-13

Query Match          98.7%; Score 1656; DB 4; Length 358;
Best Local Similarity 98.5%; Pred. No. 3.8e-150;
Matches 330; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGOIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 60
Db 7 AIDENKOKALAAALGOIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126

Qy 121 TGEQALBICDALARSGAVDIVVDSVAALTPKABIEIEIGDGHMGLAARMMSQAMRKLKAG 180
Db 127 TGEQALBICDALARSGAVDIVVDSVAALTPKABIEIEIGDGHMGLAARMMSQAMRKLKAG 186

Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246

Qy 241 ETRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLEKAGAWYSYKGEKIQ 300
Db 247 ETRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLEKAGAWYSYKGEKIQ 306

Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 11
US-10-194-686-11
; Sequence 11, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE

```

```
; TITLE OF INVENTION: SEQUENCE RECOMBINATION
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-194-686-11

Query Match      98.6%; Score 1655; DB 4; Length 358;
Best Local Similarity 98.8%; Pred. No. 4.8e-150;
Matches 331; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 126
Qy 121 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLAG 180
Db 127 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLAG 186
Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
Qy 241 ETRVVKVKNKIAAPFKQAEFQILYEGEINFGELVDLGVKEKLEKAGAWTSYKGEKIGQ 300
Db 247 ETRVVKVKNKIAAPFKQAEFQILYEGEINFGELVDLGVKEKLEKAGAWTSYKGEKIGQ 306
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANAALWLNKGNPETAKEIEKKVRELLLSNPSTP 341

RESULT 12
US-10-282-122A-75948
; Sequence 75948, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELTRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
```

```
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75948
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Salmonella typhi
; US-10-282-122A-75948

Query Match      97.6%; Score 1637; DB 4; Length 353;
Best Local Similarity 97.0%; Pred. No. 2.5e-148;
Matches 325; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 61
Qy 61 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 121
Qy 121 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLAG 180
Db 122 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLAG 181
Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 182 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 241
Qy 241 ETRVVKVKNKIAAPFKQAEFQILYEGEINFGELVDLGVKEKLEKAGAWTSYKGEKIGQ 300
Db 242 ETRVVKVKNKIAAPFKQAEFQILYEGEINFGELVDLGVKEKLEKAGAWTSYKGEKIGQ 301
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 302 GKANATTWLNKGNPETAKEIEKKVRELLLSNPSTP 336
```

```
RESULT 13
US-10-282-122A-59525
; Sequence 59525, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELTRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
```

FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
PRIOR FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 59525
LENGTH: 352
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-10-282-122A-59525

Query Match 95.9%; Score 1609; DB 4; Length 352;
Best Local Similarity 95.2%; Pred. No. 1.2e-145;
Matches 319; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGSDRSMDVKTISTGSLSDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGSDRSMDVKTISTGSLSDIALGAGGLPMGR 61
Qy 61 IVEIYGESSGKTTTLQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGESSGKTTTLQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 121
Qy 121 TGEQALICDALARSGADVIVVDSVAALTPKAEIEGEGISHMGLAARMMSQAWRKLKAG 180
Db 122 TGEQALICDALARSGADVIVVDSVAALTPKAEIEGEGISHMGLAARMMSQAWRKLKAG 181
Qy 181 NLKQSNLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
Db 182 NLKQSNLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 241
Qy 241 ETRVKVVKNTIAAPFKQAEFOILYGEINFGYELVDLGVEKLEKAGAWTSYKGEKIQG 300
Db 242 ETRVKVVKNTIAAPFKQAEFOILYGEINFGYELVDLGVEKLEKAGAWTSYKGEKIQG 301
Qy 301 GKANATAWLKDNPTAKEIEKKVRELLLSNPSTP 335
Db 302 GKANAITWLNKPNPAAKEIEKKVRELLLNQDQKP 336

RESULT 14
US-10-282-122A-55632
Sequence 55632, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
PRIOR FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 55632
LENGTH: 352
TYPE: PRT
ORGANISM: Enterobacter cloacae
US-10-282-122A-55632

Query Match 94.7%; Score 1589; DB 4; Length 352;
Best Local Similarity 94.3%; Pred. No. 1e-143;
Matches 316; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGSDRSMDVKTISTGSLSDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGSDRSMDVKTISTGSLSDIALGAGGLPMGR 61
Qy 61 IVEIYGESSGKTTTLQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGESSGKTTTLQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 121
Qy 121 TGEQALICDALARSGADVIVVDSVAALTPKAEIEGEGISHMGLAARMMSQAWRKLKAG 180
Db 122 TGEQALICDALARSGADVIVVDSVAALTPKAEIEGEGISHMGLAARMMSQAWRKLKAG 181
Qy 181 NLKQSNLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
Db 182 NLKQSNLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 241
Qy 241 ETRVKVVKNTIAAPFKQAEFOILYGEINFGYELVDLGVEKLEKAGAWTSYKGEKIQG 300
Db 242 ETRVKVVKNTIAAPFKQAEFOILYGEINFGYELVDLGVEKLEKAGAWTSYKGEKIQG 301
Qy 301 GKANATAWLKDNPTAKEIEKKVRELLLSNPSTP 335
Db 302 GKANAISWLNKPNPAAKEIEKKVRELLLNQDQSKP 336

RESULT 15
US-10-282-122A-78524
Sequence 78524, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel

APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 78524
LENGTH: 356
TYPE: PRT
ORGANISM: Versinia pectis
US-10-282-122A-78524

Query Match 91.2%; Score 1530; DB 4; Length 356;
Best Local Similarity 91.6%; Pred. No. 4.8e-138;
Matches 304; Conservative 15; Mismatches 13; Indels 0; Gaps 0;

Qy	1	AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVKTISTGSLSLDIALGAGGLPMGR	60
Db	2	AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVKTISTGSLSLDIALGAGGLPMGR	61
Qy	61	IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD	120
Db	62	IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD	121
Qy	121	TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG	180
Db	122	TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG	181
Qy	181	NLKQSNLTLLIFNQIRMKIGVMFGNPGPTTGGNALKFYASVRLDIRRIGAVKEGENVYGS	240
Db	182	NLKQSNLTLLIFNQIRMKIGVMFGNPGPTTGGNALKFYASVRLDIRRIGAVKEGENVYGS	241
Qy	241	ETRVKVNKNIAAPFKQAEFQILYEGINFGELVGLVKEKLEKAGAWYSYKGEKIGQ	300
Db	242	ETRVKVNKNIAAPFKQAEFQILYEGINFGELVGLVKEKLEKAGAWYSYKGEKIGQ	301
Qy	301	GKANATAWLKONPETAKEIEKKVRELLSNPN	332
Db	302	GKANASNVLKENPAIAAELDKKREMLNGGN	333

Search completed: February 16, 2006, 01:44:35
Job time : 124 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 01:40:35 ; Search time 10.5 Seconds
(without alignments)
453.421 Million cell updates/sec

Title: US-10-733-782-3

Perfect score: 1678
Sequence: 1 AIDENKOKALAAALGQIEKQ.....AKEIEKVKRELLSNPNSTP 335

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 107799 seqs, 14211699 residues

Total number of hits satisfying chosen parameters: 107799

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA_New*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1208.5	72.0	375	6	US-10-467-657-4766 Sequence 4766, Ap
2	1019	60.7	363	7	US-11-074-176-296 Sequence 296, App
3	970.5	57.8	355	7	US-11-098-686-11427 Sequence 11427, A
4	154	9.2	443	7	US-11-098-686-10861 Sequence 10861, A
5	111.5	6.6	459	6	US-10-467-657-3092 Sequence 3092, Ap
6	109	6.5	453	6	US-10-467-657-206 Sequence 206, App
7	109	6.5	453	6	US-10-467-657-6400 Sequence 6400, Ap
8	104	6.2	454	7	US-11-098-686-10599 Sequence 10599, A
9	102	6.1	453	6	US-10-467-657-3626 Sequence 3626, Ap
10	100	6.0	925	6	US-10-454-437-50 Sequence 50, Appl
11	99	5.9	291	6	US-10-467-657-4148 Sequence 4148, Ap
12	97	5.8	376	7	US-11-213-368-14 Sequence 14, Appl
13	96	5.7	286	7	US-11-082-389-88 Sequence 88, Appl
14	96	5.7	471	6	US-10-467-657-802 Sequence 802, App
15	95	5.7	257	7	US-11-156-084-264 Sequence 264, App
16	95	5.7	376	7	US-11-082-389-122 Sequence 122, App
17	95	5.7	376	7	US-11-082-389-124 Sequence 124, App
18	94.5	5.6	548	6	US-10-793-626-326 Sequence 326, App
19	94.5	5.6	640	7	US-11-205-109-9 Sequence 9, Appl
20	94.5	5.6	716	6	US-10-131-826A-96 Sequence 96, Appl
21	94	5.6	365	7	US-11-082-389-132 Sequence 132, App
22	94	5.6	405	6	US-10-467-657-7420 Sequence 7420, Ap
23	94	5.6	479	7	US-11-082-389-130 Sequence 130, App
24	94	5.6	655	7	US-11-094-586-10 Sequence 10, Appl
25	92.5	5.5	423	7	US-11-070-080-16 Sequence 16, Appl

26	92	5.5	302	6	US-10-793-626-2798 Sequence 2798, Ap
27	92	5.5	376	7	US-11-213-368-6 Sequence 6, Appl
28	92	5.5	650	6	US-10-878-556A-110 Sequence 110, App
29	92	5.5	806	7	US-11-108-172-1117 Sequence 1117, Ap
30	92	5.5	806	7	US-11-108-172-1118 Sequence 1118, Ap
31	92	5.5	806	7	US-11-079-900-1 Sequence 1, Appl
32	91.5	5.5	419	6	US-10-979-821-6 Sequence 6, Appl
33	91.5	5.5	419	7	US-11-114-922-6 Sequence 6, Appl
34	91.5	5.5	569	6	US-10-506-443A-35 Sequence 35, Appl
35	90.5	5.4	718	7	US-11-074-176-306 Sequence 306, App
36	90.5	5.4	723	7	US-11-074-176-18 Sequence 18, Appl
37	89.5	5.3	690	7	US-11-212-443-171 Sequence 171, App
38	89.5	5.3	710	7	US-11-212-443-30 Sequence 30, Appl
39	89	5.3	885	6	US-10-467-657-2302 Sequence 2302, Ap
40	89	5.3	1992	7	US-11-013-759-3 Sequence 3, Appl
41	89	5.3	1992	7	US-11-013-759-13 Sequence 13, Appl
42	89	5.3	2047	7	US-11-013-759-4 Sequence 4, Appl
43	89	5.3	2047	7	US-11-013-759-7 Sequence 7, Appl
44	88.5	5.3	687	7	US-11-212-443-173 Sequence 173, App
45	88.5	5.3	710	7	US-11-212-443-28 Sequence 28, Appl

ALIGNMENTS

RESULT 1

US-10-467-657-4766
; Sequence 4766, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10467.657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 4766
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4766

Query Match		72.0%	Score 1208.5;	DB 6;	Length 375;
Best Local Similarity		68.7%	Pred. No. 5.8e-86;		
Matches 224;		Conservative 53;	Mismatches 48;	Indels 1;	Gaps 1;
Qy	2	IDENKOKALAAALGQIEKQFGKGSIMRL-GEDRSMVKTITSTGSLSLDIALGAGGLPMGR	60		
Db	28	MSDDKSKALAAALQAEKSPKAIMKMDGSGQBEENLEVIETGSLGLDLALGVGGLPRGR	87		
Qy	61	IVFIYGPSSGKTTLTLOVIAAAREKTCAPDAEHALDPIYARKLGVDIDNLLCSQPD	120		
Db	88	IVFIYGPSSGKTTLCLEAVACQCKNGGVCAFDVAEHAFFVYARKLGKVBELYSQPD	147		
Qy	121	TGQALIEICALARSGAVDVIIVVDSVAALTPKABIEGIEIGDSHMGLAARMWSQAMRKLKAG	180		
Db	148	TGQALIEICUTLVRSBGIDVWVDSVAALVPKAIEGDMGDSHVGLQARLMSQALRKLTG	207		
Qy	181	NLKQSNLLIFINQIRMKIGVMFGNPTTTGGNALKFPYASVRLDIRRIGAVKEGENVVGVS	240		
Db	208	HIKKTNTLVVFIQIRMKIGVMFGSPETTTGGNALKFPYSSVRLDIRRTGSIKKGEVLGN	267		
Qy	241	ETRVKVKNKIAAPFKQAEFOILYGGGINFYGVLDLGVKEKLEKAGAWSYKGEKIGQ	300		
Db	268	ETRVKVKNKVAPPFPQAEFDILYBEGISWEGELIDIVGNKNDIINKSGAWSYNGAKIGQ	327		

```
Qy 301 GKANATAWLKNDPETAKEIEKKVREL 326
Db 328 GKONRVWLXENPEISDEIDAKIRAL 353

RESULT 2
US-11-074-176-296
; Sequence 296, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Perill, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; Stresses-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 296
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-296

Query Match 60.7%; Score 1019; DB 7; Length 363;
Best Local Similarity 56.8%; Pred. No. 2.2e-71;
Matches 183; Conservative 74; Mismatches 65; Indels 0; Gaps 0;

Qy 4 ENKOKALAAALGOIEKQFGKSGIMRLGEDRSDMDVKTISTGSLSDIALGAGGLPMGRIVE 63
Db 4 DEKKAALDRAALKKIEKNGKGAVMRGEKATQISTVTGSLDALDAIGVGGYPRGRIIE 63

Qy 64 IYGPSSGKTTLTLOVIAAAQREGKTCAPIDAHAHALDPIYARKLGVDIDNLLCSQPDGTG 123
Db 64 VYGPSSGKTTVALHAVAQVKGRTAAVYDAENAMDPAAYEAALGVDIDSLISQPDGTG 123

Qy 124 QALEICDALARGAVDVVDSVAALTPKAEIEGIDSHMGLAARMMSQAMRKLAGNLK 183
Db 124 EGLQADTLISSGADIDVVDVVAALVPRAEIEGEMGDHVGLOQLRLMSQALRLKLSGTIS 183

Qy 184 OSNTLLIFINQIRMKIGVFGNPETTTGCGNALKFYASVRLDIRRIGAVKEGENVVGSETR 243
Db 184 KTKTIAIFINQIREKVGWFGNPETTPGGRALKFYSTVRLEVRRAEQIKQSGDVLGNRVK 243

Qy 244 VVKVKNKIAAPPKQAEFOILYGEINFGYELVDLGKVEKLEKAGAWSYKGEKIGQOKA 303
Db 244 IKVVKNKVAAPPKVAEVDIMYKGISQSGELLDMAADKDIIDKAGSWYSKSDRIGQGRE 303

Qy 304 NATAWLKNDPETAKEIEKKVRE 325
Db 304 NAKKYLEEHPDIYQVQSQVRQ 325

RESULT 3
US-11-098-686-11427
; Sequence 11427, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10861
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-11427

Query Match 9.2%; Score 154; DB 7; Length 443;
Best Local Similarity 32.1%; Pred. No. 0.00011;
Matches 62; Conservative 28; Mismatches 81; Indels 22; Gaps 8;

Qy 18 EKQKGSIMRLGEDRSDMDVKTISTGSLSDIALGAGGLPMGRIVEIYGPSSGKTTTL 77
Db 44 QKTYKUKPIPIGRVETSTYQPFSTGKLVLDHILGKLVPSALL-IGGEFGIGKSTLLL 102

Qy 78 QVIAAAREGKTCAPIDAHAHALDPIYAR--KLGVDIDNLLCSQPDGTGQALEICDALRS 135
Db 103 QLAGATATLGKIVLYVSGEESLPQIKARAERLTVLHNDLLSM--ATSQVEDILPLNTS 159
```

```
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11427
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-11427

Query Match 57.8%; Score 970.5; DB 7; Length 355;
Best Local Similarity 57.3%; Pred. No. 1.1e-67;
Matches 189; Conservative 64; Mismatches 76; Indels 1; Gaps 1;

Qy 3 DENKOKALAAALGOIEKQFGKSGIMRLGEDRSDMDVKTISTGSLSDIALGAGGLPMGRIV 62
Db 9 EDSRRALKATLDTIERKFGQGAVMKLSDDVVHVAVIPTSGISGLDLAGLGGIPRGRVT 68

Qy 63 EYGPSSGKTTLTLOVIAAAQREGKTCAPIDAHAHALDPIYARKLGVDIDNLLCSQPDGTG 122
Db 69 EYGPSSGKTTLTTHIAECQKLGTTAAFDIAHAHALDIAYAKRLGVKTDLEIISQPDHG 128

Qy 123 EQALEICDALARGAVDVVDSVAALTPKAEIEGIDSHMGLAARMMSQAMRKLAGNL 182
Db 129 EQALEIADMLVRSQAVDLVVDSVAALIPQTELESGMGETQVGGHARLMSHALRKLGTI 188

Qy 183 KQSNLTLLIFINQIRMKIGVM-FGNPETTTGCGNALKFYASVRLDIRRIGAVKEGENVVGE 241
Db 189 HKSHTAVIFINQIRMKIGVVGYPETTTGCGNALKFYSSVNRMDIRKIQTLLKDKBESYSSL 248

Qy 242 TRVKVKNKIAAPPKQAEFOILYGEINFGYELVDLGKVEKLEKAGAWSYKGEKIGQG 301
Db 249 TRVKVKNKVAAPPKPAKFDIINGTGISRSGELIDLGVEAGIIDKSGSWFAFGSEKLGQG 308

Qy 302 KANATAWLKNDPETAKEIEKKVRELLSNP 331
Db 309 KEKVRALLDENTSLLKSSIIETALMEHLGNP 338
```

```
RESULT 4
US-11-098-686-10861
; Sequence 10861, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10861
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10861
```

```
Query Match 9.2%; Score 154; DB 7; Length 443;
Best Local Similarity 32.1%; Pred. No. 0.00011;
Matches 62; Conservative 28; Mismatches 81; Indels 22; Gaps 8;

Qy 18 EKQKGSIMRLGEDRSDMDVKTISTGSLSDIALGAGGLPMGRIVEIYGPSSGKTTTL 77
Db 44 QKTYKUKPIPIGRVETSTYQPFSTGKLVLDHILGKLVPSALL-IGGEFGIGKSTLLL 102

Qy 78 QVIAAAREGKTCAPIDAHAHALDPIYAR--KLGVDIDNLLCSQPDGTGQALEICDALRS 135
Db 103 QLAGATATLGKIVLYVSGEESLPQIKARAERLTVLHNDLLSM--ATSQVEDILPLNTS 159
```


Qy 147 AALTPKA--EIEGEIGDSHMGAAARMMSQAMRKLAGNLKQNTLLIFINQIRMKIGVMFG 204
Db 321 -HIMPRAGRDVAELGN-----ISRRLKNLAEL-----NTPVVLVAQLNR-----G 361
Qy 205 NPETTTGKNALKFYASVR---LDIRIRIGAVKEGENVV 238
Db 362 NTK-----QADKEPNWADIRGSGAIEQDANII 388

RESULT 8
US-11-098-686-10599
; Sequence 10599, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10599
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10599

Query Match 6.2%; Score 104; DB 7; Length 454;
Best Local Similarity 22.0%; Pred. No. 0.84;
Matches 55; Conservative 45; Mismatches 94; Indels 56; Gaps 11;
Qy 16 QIEQFGKSGIMRWGEDRSMVDKTIISTGSLDIALGAGGLPMGRIVEIYGPSSGKTL 75
Db 168 ELEKRF-----ERKEQVTGTTGYNRLD-KLTAGLQPSDLIIIVAARP-SMGKTAF 215
Qy 76 TLQV-IAAAREGKTCAPIDAEHALDPIYARKL-----GVIDDNLCS--QPTDGEQALEI 128
Db 216 SLNWMRSAIQGTPTVAIYSLEMSNQLMWRMLCAWGKVDLSLHRLHGYLNSDEWSRLYHA 275
Qy 129 CDALARSQAVDIVVDSVAALTP-----KAEIEGI-----GDSHMGGLAA 168
Db 276 ADVLQQA---PIFIDTTPALSPLELRARTRLKTESDVLGWLVDYLQLMRGNKRTDSRE 331
Qy 169 RMMSQAMRKLAGNLKQNTLLIFINQIRMKIGVMFGNPETTTGKNALKFYASVRLDIRRI 228
Db 332 QEISISRLSKSLAKEINIPVALSQLNRKL-----EDRTDKRPQLS-----DLRES 378
Qy 229 GAVKEGENVV 238
Db 379 GAIEQDADVI 388

RESULT 9
US-10-467-657-3626
; Sequence 3626, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8

; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 3626
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3626

Query Match 6.1%; Score 102; DB 6; Length 453;
Best Local Similarity 23.7%; Pred. No. 1.2;
Matches 69; Conservative 43; Mismatches 85; Indels 94; Gaps 18;
Qy 2 IDENKQKA--LAAALQIEKQKQKGSIMRL-----GKDAVKRETKTFGQTVEDLIGLQKRLDGVRFGLPTG 43
Db 138 VAEKLSRAADELAAA-----GKDAVKRETKTFGQTVEDLIGLQKRLDGVRFGLPTG 189
Qy 44 SLSLDIALG-AGGLPMGRIVEIYGPSSGKTLTLQVIAAAQREGKTCAPIDAEHALDPI 102
Db 190 LMKLD---GMTGGLPDGNLIVIAARPSMGKTVLAENIARFALKQKAVHFQSYEMSAVEL 246
Qy 103 ----YARKLGVDIDNL-----LCSQPDTEQALE-----ICDAL----- 132
Db 247 ARRGMAAECPNIPMONLKTGNLTQSDYANMPITYVSAKEWKFPDNCDDLINVDLCLFLAKEK 306
Qy 133 ARSGAVDIVVDSVAALTPKA--EIEGEIGDSHMGGLAARMMSQAMRKLAGNLKQNTLLI 190
Db 307 KLTGDLVVDHL-HIMPRAGRDVAELGN-----ISRRLKNLAEL-----NTPVV 353
Qy 191 FINQIRMKIGVMFGNPETTTGKNALKFYASVR---LDIRIRIGAVKEGENVV 238
Db 354 LVAQLNR-----GNTK-----QADKEPNWADIRGSGAIEQDANII 388

RESULT 10
US-10-454-437-50
; Sequence 50, Application US/10454437
; Publication No. US20050277115A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Heberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; FILE REFERENCE: BGI-128CPCN
; CURRENT APPLICATION NUMBER: US/10/454,437
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 50
; LENGTH: 925


```
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-454-437-50

Query Match
Best Local Similarity 6.0%; Score 100; DB 6; Length 925;
Matches 75; Conservative 67; Mismatches 125; Indels 120; Gaps 18;

Qy 14 LGQIEKQFGKGS--IMRLGED-----RSMVDVTKIS-----TGSLSLDIALGAG 54
Db 110 LGLIREGEGVAAQVLVLKGLADLPRVRQVQIQLLSGVEGGSGSPGGQGAPTGDDAVGAG 169

Qy 55 GLPMGRIVEIYGPSSG-----KTYLTL-----QVIAAARQEGKTCAFIDAEHALDPIYAR 105
Db 170 AAPGGR-----PSSGSGFGERSTSLVLDQFGRNLTAQADGK-----LDPVVGR 212

Qy 106 KLGVD-----IDNLCSOPDTGEQALEICDALARSGAVDVIIVDSVAALTPKAE 154
Db 213 DKEIERIMQVLSRRTKNNPVLIGEPGVGTAV-----VEGLALDIVNGKVPETL 261

Qy 155 IEGEIGDSHMG--LAARM-----MSQAMRKLAKNLSKQNTLLIFINQIRMKIGVMFGNPET 208
Db 262 KDKQVSYLDLGLSLVAGSYRGDFEERLKKVLKEINQRGDIILFIDEIHTLVGA--GAERG 319

Qy 209 TTGGNALKFYASVRLDIRIGAV-----KGENNVVSGSETRVKVVKNKIAAPFKQAEFQI 262
Db 320 AIDAASLLKPKLARGELQIGATTILDEYRKHIEKDAALERFQPV--QVPEPSVDLTVEI 377

Qy 263 LYG-----EGINFYCELVDLGVKEK-----LIEKAGANYSYKGEKIG 299
Db 378 LKGLRDYEAHHRVITDGAUATAAQADRVINDRFLPDKAVALIDEAGARMRIK----- 432

Qy 300 QGKANATAWLKNDPETAKEIEKKVREL 326
Db 433 -----RWTA-----PSSLREVDERTADV 450

RESULT 11
US-10-467-657-4148
; Sequence 4148, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 4148
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4148

Query Match
Best Local Similarity 5.9%; Score 99; DB 6; Length 291;
Matches 63; Conservative 44; Mismatches 83; Indels 76; Gaps 15;

Qy 55 GLPM---GRIVEIYGPSSGKTTTLQVIAAARQEGKTCAFID--AEHALDPIYARKLGV 109
Db 70 GIPLNNRQKVIAIDPGASGKGTVAARKVAALGYD-----YLDTGALYRLTALYAKQGV 124

Qy 110 ---DIDNLCSOPTGQALEICDALAR-----SGAVDVIIVDSVAALTPKABIEGIG 160
Db 125 EWHDEENV-----SALAKGLPAVFESG--NRILLDG-----EDVSDGIR 160

; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-454-437-50

Query Match
Best Local Similarity 5.8%; Score 97; DB 7; Length 376;
Matches 53; Conservative 30; Mismatches 80; Indels 62; Gaps 11;

Qy 58 MGRIVEIY--GPSSGKTTTLTLOVIAAARQEGKTCAFIDAEHALDPI-YARKLGV--ID 112
Db 46 MPILLRVYIDGPHCMGKTTTQLLVALGSRD-----DIVYVDPPTYWQVLASSETIA 98

Qy 113 NLLCSQP-----DTGEQALEICDALARSGAVDVIIVDSVAALTPKABIEGIGDSHM- 164
Db 99 NIVTTQHRLDQGEISAGDAAVVMTSAQITWGM-PPYAVTDAVLA-----PHIGGEAGSSHAP 153

Qy 165 -----GLAARMMSQAMRKLAKNLSKQNTLLIFINQIRMKIGVMFGNPETTTGCGN 213
Db 154 PPALTILFDRHRPIAALCYPAARYLMGSM--TPQAVLAFVALI-----PPTLPFTN 202

Qy 214 -----ALKFYASVRLDIRRIGAVKGENVVGSETR 243
Db 203 IVGALPEDRHIDRLAKRQRPGERLDLAMLAIIRRVYGLLANTVR 247

RESULT 13
US-11-082-389-88
; Sequence 88, Application US/11082389
; Publication No. US20050244935A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORNYNEACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
```


Db 196 YFANQYDLSEVETLIKRDSSRRYAKKQLTYFXNQFPDMRWFDTKQITENPKLIIDLVKKFN 255
QY 325 E 325
Db 256 Q 256

Search completed: February 16, 2006, 01:45:02
Job time : 11.5 secs

This Page Blank (uspto)